

NUMBER OF SEQ ID NOS: 10535
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9696
LENGTH: 2494
TYPE: DNA
ORGANISM: Homo sapiens
US-09-649-163-9696

Query Match 98.5% Score 1548.8 DB 25 Length 2494
Best Local Similarity 99.5% Pred. No. 2.9e-298
Matches 1564: Conservative 0: Mismatches 7: Indels 1: Gaps 1:

```
QY 1 GCCCCACAGCCGCGCCCTGGCAGCCCGCTGGGACACCGATTAAGAGCTGAAGCAG 60
DB 882 GCCCCACAGCCGCGCCCTGGCAGCCCGCTGGGACACCGATTAAGAGCTGAAGCAG 941
QY 61 ACCCCGCGCCAGCGGAGCG -CCCCACAGCCCGCAGGAGCCCTGGCAGCGGAGCCGCG 119
DB 942 ACCCCGCGCCAGCGGAGCGGCCCCACAGCGCCAGGAGCCCTGGCAGCGGAGCGCG 1001
QY 120 GGTGAGGTTATGATCCAGCGGCGGCGCCGCGGCGGTGCTCCCGCGCGCGCGGT 179
DB 1002 GGTGAGGTTATGATCCAGCGGCGGCGCCCGCGGCGGTGCTCCCGCGCGCGCGGT 1061
QY 180 GGTGAGGTTATGATCCAGCGGCGGCGGCAAGGCGCAAGGCTTGGCAGCTTCCGAGTCA 239
DB 1062 GGTGAGGTTATGATCCAGCGGCGGCGGCAAGGCGCAAGGCTTGGCAGCTTCCGAGTCA 1121
QY 240 GGTGAGGTTATGATCCAGCGGCGGCGGCAAGGCTTGGCAGCTTCCGAGTCA 239
DB 1122 GGTGAGGTTATGATCCAGCGGCGGCGGCAAGGCTTGGCAGCTTCCGAGTCA 1181
QY 300 GAACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 359
DB 1182 GAACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1241
QY 360 CATGCTGAGAGAGGAGCTGATGACAGAGGTGATGAGAGGAGCTTATGAGCGGCGCTACTG 419
DB 1242 CATGCTGAGAGAGGAGCTGATGACAGAGGTGATGAGAGGAGCTTATGAGCGGCGCTACTG 1301
QY 420 GAGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 479
DB 1302 GAGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1361
QY 480 AGCTTCCTTGAACATTAATGCTGCTGATGAGAGGCTGACCAATGAAGAGCTTCTGACCA 539
DB 1362 AGCTTCCTTGAACATTAATGCTGCTGATGAGAGGCTGACCAATGAAGAGCTTCTGACCA 1421
QY 540 CTGACAGCTATTGCTGCGCGCGGCGGCTGCTGACCAATGAAGAGCTTCTGACCA 599
DB 1422 CTGACAGCTATTGCTGCGCGCGGCGGCTGCTGACCAATGAAGAGCTTCTGACCA 1481
QY 600 GCGTTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 659
DB 1482 GCGTTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1541
QY 660 GAGCTAGAGAGAGAGATGAGAGGCTGCGGCGGAGATGCGCTTACTCTGAGGAGCTT 719
DB 1542 GAGCTAGAGAGAGAGATGAGAGGCTGCGGCGGAGATGCGCTTACTCTGAGGAGCTT 1601
QY 720 CCGTGCCTTGGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 779
DB 1602 CCGTGCCTTGGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1661
QY 780 AGTGGCTTCCAGAGACAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 839
DB 1662 AGTGGCTTCCAGAGACAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1721
QY 840 CCGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 899
DB 1722 CCGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1781
QY 900 TGTGCTAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 959
```

```
DB 1782 TGTGCTAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1841
QY 960 GGGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1019
DB 1842 GGGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1901
QY 1020 CATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1079
DB 1902 CATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1961
QY 1080 CTACTTGTATATGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1139
DB 1962 CTACTTGTATATGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2021
QY 1140 GTTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1199
DB 2022 GTTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2081
QY 1200 CTACTTGTATATGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1259
DB 2082 CTACTTGTATATGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2141
QY 1260 GCCACCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1319
DB 2142 GCCACCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2201
QY 1320 AGGAGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1379
DB 2202 AGGAGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2261
QY 1380 GAGACTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1439
DB 2262 GAGACTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2321
QY 1440 TGAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1499
DB 2322 TGAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2481
QY 1500 TTTGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1559
DB 2382 TTTGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2441
QY 1560 AAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1571
DB 2442 CACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2453
```

RESULT 9
US-09-652-126-10026
Sequence 10026, Application US/09652126
GENERAL INFORMATION:
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: 1600, 1185-001
CURRENT APPLICATION NUMBER: US/09/652,126
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/151,132
NUMBER OF SEQ ID NOS: 10051
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10026
LENGTH: 2494
TYPE: DNA
ORGANISM: Homo sapiens
US-09-652-126-10026

Query Match 98.5% Score 1548.8 DB 25 Length 2494:
Best Local Similarity 99.5% Pred. No. 2.9e-298
Matches 1564: Conservative 0: Mismatches 7: Indels 1: Gaps 1:

```
QY 1 GCCCCACAGCCGCGCCCTGGCAGCCCGCTGGGACACCGATTAAGAGCTGAAGCAG 60
```



```
Db 1242 CATGCTGAGAGAGGCTGTATGACAGAGGTGTGAAGGGCTCATGAGCGGCTGACGTG 1301
Oy 420 GGAGACCGGCATTCAGAGAGCCCTGTGTAGCTTCCAGCAGAGCTTGGCAAGCGGCTGAC 479
Db 1302 GGAGACCGGCATTCAGAGAGCCCTGTGTAGCTTCCAGCAGAGCTTGGCAAGCGGCTGAC 1361
Oy 480 AGCTTCCTTGAACATTTATGCTGCTATGAGCAGAGTACCAATGAAGACCTCTGACCAC 539
Db 1362 AGCTTCCTTGAACATTTATGCTGCTATGAGCAGAGTACCAATGAAGACCTCTGACCAC 1421
Oy 540 CTGACAGCTATTGCTGTGGCGCGGCTGTGTACCAAGCAATGCTGTGTGTGTGTGTGTGT 599
Db 1422 CTGACAGCTATTGCTGTGGCGCGGCTGTGTACCAAGCAATGCTGTGTGTGTGTGTGTGT 1481
Oy 600 GGGCTGGGGGCTGGCGCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 639
Db 1482 GGGCTGGGGGCTGGCGCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1541
Oy 660 GGACCTTAGAGAGTGAAGAGTATCGGCGCTGTGGGGAGATGCCCTTCACTCTGGGACCTT 719
Db 1542 GGACCTTAGAGAGTGAAGAGTATCGGCGCTGTGGGGAGATGCCCTTCACTCTGGGACCTT 1601
Oy 720 CTTGCGCTTGGAGCGCTTGTGCGACCTACCGCGGCGCACTGGGCTACCTTCCCTGTAGAG 779
Db 1602 CTTGCGCTTGGAGCGCTTGTGCGACCTACCGCGGCGCACTGGGCTACCTTCCCTGTAGAG 1661
Oy 780 AGTGGGTTCCAGAGACAGCTGCTCCCTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 839
Db 1662 AGTGGGTTCCAGAGACAGCTGCTCCCTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1721
Oy 840 CCTGTGCTCCAGTGAAGAGAGCAGTGTGCTCTCACTGTGACAGTGTGTGTGTGTGTGTGT 899
Db 1722 CCTGTGCTCCAGTGAAGAGAGCAGTGTGCTCTCACTGTGACAGTGTGTGTGTGTGTGTGT 1781
Oy 900 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 959
Db 1782 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1841
Oy 960 GGGCGGCTGTGAGAGCGGCGGTATGACATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1019
Db 1842 GGGCGGCTGTGAGAGCGGCGGTATGACATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1901
Oy 1020 CATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1079
Db 1902 CATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1961
Oy 1080 CTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1139
Db 1962 CTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2021
Oy 1140 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1199
Db 2022 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2081
Oy 1200 CTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1259
Db 2082 CTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2141
Oy 1260 GCCACCGGCGAAGAGCGCTTATGACCGCTGTGGCGCGGCTGTGTGTGTGTGTGTGTGTGT 1319
Db 2142 GCCACCGGCGAAGAGCGCTTATGACCGCTGTGGCGCGGCTGTGTGTGTGTGTGTGTGTGT 2201
Oy 1320 AGGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1379
Db 2202 AGGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2261
Oy 1380 GAGACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1439
Db 2262 GAGACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2321
Oy 1440 TGAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1499
Db 2322 TGAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2381
```

```
Oy 1500 TTTGTGTGTGAGCCCGCCAGACCAATCCAAATGAAGTGTATGATTCAGGCTGCA 1559
Db 2382 TTTGTGTGTGAGCCCGCCAGACCAATCCAAATGAAGTGTATGATTCAGGCTGCA 2441
Oy 1560 AAAAAAAAAAAAAA 1571
Db 2442 CACAAACATGAA 2453

RESULT 12
US-09-663-784-702
? Sequence 702, Application US/09663784
? GENERAL INFORMATION:
? APPLICANT: GALVIN, Katherine
? TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES THEREFOR
? FILE REFERENCE: 1600,1160-001
? CURRENT APPLICATION NUMBER: US/09/663,784
? PRIOR FILING DATE: 2000-09-18
? PRIOR APPLICATION NUMBER: 60/154,571
? NUMBER OF SEQ ID NOS: 830
? SEQ ID NO 702
? SOFTWARE: FASTSEQ for windows version 4.0
? LENGTH: 2494
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-663-784-702

Query Match 98.5%; Score 1548.8; DB 26; Length 2494;
Best Local Similarity 99.5%; Pred. No.2,9e+298;
Matches 1564; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Oy 1 GCCCACAAGCGCGCTGTGAGAGCGCGGCTGTGAGAGCAACGATGAAGCTGAAGACAG 60
Db 882 GCCCACAAGCGCGCTGTGAGAGCGCGGCTGTGAGAGCAACGATGAAGCTGAAGACAG 941
Oy 61 AGCGCGCGCGCACGGGACAGC-CCGACAGCGCGGACAGCGCCCTGTGTGTGTGTGTGTGT 119
Db 942 AGCGCGCGCGCACGGGACAGC-CCGACAGCGCGGACAGCGCCCTGTGTGTGTGTGTGTGT 1001
Oy 120 GGTGAGGTTATGATCCAGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
Db 1002 GGTGAGGTTATGATCCAGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1061
Oy 180 GGTGTGTGTGTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
Db 1062 GGTGTGTGTGTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1121
Oy 240 CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 299
Db 1122 CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1181
Oy 300 GAACCAAGCGCGGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 359
Db 1182 GAACCAAGCGCGGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1241
Oy 360 CATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 419
Db 1242 CATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1301
Oy 420 GGAAGACCGGCATCCAGAAAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 479
Db 1302 GGAAGACCGGCATCCAGAAAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1361
Oy 480 AGCTTCTTTGAACCATTTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 539
Db 1362 AGCTTCTTTGAACCATTTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1421
Oy 540 CTGACAGCTATTGCTGTGGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db 1422 CTGACAGCTATTGCTGTGGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1481
```


QY	600	GGCTTTGGGGCTGGGCTCTTCTCTGTGTGTCAGACCTGGGCTGGGCTTCATTGGCGATGT	659
Db	1482	GGCTTCGGGGCTGGGCTCTTCTCTGTGTGTCAGACCTGGGCTGGGCTTCATTGCTGATGT	1541
QY	660	GGACCTAGAGAGTGAAGATGAGGGCTGTGGGGGAGATGAGCTTCACTCTGGGACCTT	719
Db	1542	GGACCTAGAGAGTGAAGATGATGSGGCTGTGGGGGAGATGSGCTTCACTCTGGGACCTT	1601
QY	720	CCTGCGTCTGGCAGCCCTGGGACCTACCGGGGGGAGCTGGCTTACCTCCTGTAGAGAG	779
Db	1602	CCTGCGTCTGGCAGCCCTGGGACCTACCGGGGGGAGCTGGCTTACCTCCTGTAGAGAG	1661
QY	780	AGTGGGTTCGAAGACACCTTCCTCCCCGCTGTGGTGTCCAGCAGGGGCCGGTAGATGCACA	839
Db	1682	AGTGGGTTCGAAGACACCTTCCTCCCCGCTGTGGTGTCCAGCAGGGGCCGGTAGATGCACA	1721
QY	840	CCCTGTGCCACATGGAGGAGGACAGTGGCCCTCTCACTGAGACAGTGTGTGCCGAGAGGACTT	899
Db	1722	CCCTGTGCCACATGGAGGAGGACAGTGGCCCTCTCACTGAGACAGTGTGTGCCGAGAGGACTT	1781
QY	900	TGTGCTAGTCTGTGGCAGCTGCTGCACATCGCACCTGGGCAATGAGATGTTTGTGCACCCAT	959
Db	1782	TGTGCTAGTCTGTGGCAGCTGCTGCACATCGCACCTGGGCAATGAGATGTTTGTGCACCCAT	1841
QY	960	GGGCGCGGTGACACCTGGGCTCATATGCATGTTTCTACGTGGGGGGGAGTGTCTCGTGC	1019
Db	1842	GGGCGCGGTGACACCTGGGCTCATATGCATGTTTCTACGTGGGGGGGAGTGTCTCGTGC	1901
QY	1020	CATGCTGTGCGCCTCTTCTGTGGCCATGAGAGAGGGGACGACATATGAGATGAAATGCC	1079
Db	1902	CATGCTGTGCGCCTCTTCTGTGGCCATGAGAGAGGGGACGACATATGAGATGAAATGCC	1961
QY	1080	CTACTGGGTATATGTGGCCGTGTGGCTTCCGCTTGGAGGCCAAGAGATGGGAAGGTGT	1139
Db	1962	CTACTGGGTATATGTGGCCGTGTGGCTTCCGCTTGGAGGCCAAGAGATGGGAAGGTGT	2021
QY	1140	GTTTGGCAGTGGATGGGAATTGATGTTTAGCGAGGCCGTGCAGGGGACAGTGCACACAAA	1199
Db	2022	GTTTGGCAGTGGATGGGAATTGATGTTTAGCGAGGCCGTGCAGGGGACAGTGCACACAAA	2081
QY	1200	CTACTTCTGATGTGTACGCGGTTCGTGTGAGCCCGCCAGCTGGAGGCCCCAGCAGAT	1235
Db	2082	CTACTTCTGATGTGTACGCGGTTCGTGTGAGCCCGCCAGCTGGAGGCCCCAGCAGAT	2141
QY	1260	GGCACCGGCACAAGAGCCCTTATGACCCGTGGGCGCGCGTGTGTCCTTATGTACTTGTG	1319
Db	2142	GGCACCGGCACAAGAGCCCTTATGACCCGTGGGCGCGCGTGTGTCCTTATGTACTTGTG	2201
QY	1320	AGGACCCCTTCTCTCTCCCTTAGAGGCTTCAGAGGCGCTTCCACAGCTTCGTGGGGGTGAG	1379
Db	2202	AGGACCCCTTCTCTCTCTCCCTTAGAGGCGTTCAGAGGCGCTTCCACAGCTTCGTGGGGGTGAG	2261
QY	1380	GAGACTCTCTTGAGAAAGGTGAGAGGTTGAGAGCTATGCTTTGGGGGACACAGCCAGAA	1435
Db	2262	GAGACTCTCTTGAGAAAGGTGAGAGGTTGAGAGCTATGCTTTGGGGGACACAGCCAGAA	2321
QY	1440	TGAAGTCTCTGGGTAGAGAGGCCAGCTGGTGGGCCAGCGCTATGTAAAGCCTTTTAG	1499
Db	2322	TGAAGTCTCTGGGTAGAGAGGCCAGCTGGTGGGCCAGCGCTATGTAAAGCCTTTTAG	2381
QY	1500	TTTTGTTCTGAGAGCCCGACCCAGAACCAATTCAAATTAAGTGACATTTCCACAGCTGA	1555
Db	2382	TTTTGTTCTGAGAGCCCGACCCAGAACCAATTCAAATTAAGTGACATTTCCACAGCCGC	2441
QY	1560	AAAAAAAAAAAA 1571	
Db	2442	CACAAACATGAA 2453	

RESULT 13
US-09-699-998-8671
; Sequence 8671, Application US/09699998
; GENERAL INFORMATION:

```

: APPLICANT: Holtzman, Douglas A.
: APPLICANT: Gearing, David P.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: 1600.2008-001
: CURRENT APPLICATION NUMBER: US/09/699,998
: CURRENT FILING DATE: 2000-10-30
: PRIOR APPLICATION NUMBER: 60/162,362
: PRIOR FILING DATE: 1999-10-29
: NUMBER OF SEQ ID NOS: 10905
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8671
: LENGTH: 2494
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-699-998-8671

Query Match          98.5%;   Score 1548.8;   DB 27;   Length 2494;
Best Local Similarity 99.5%;   Pred. No. 2,9e+298;
Matches 1564;   Conservative 0;   Mismatches 7;   Indels 1;   Gaps 1

```

OY	1	GCCTCACAGCCGGCCCTGCGCAGCCGCCGCTGGGGACACCGTATTAGSAGAGCTGAAGGCAGG	60
Db	882	GCCTCACAGCCGGCCCTGCGCAGCCGCCGCTGGGGACACCGTATTAGSAGAGCTGAAGGCAGG	941
OY	61	AGCCCGCCGACAGGGGACGCG - CCCCACAGCGCCAGGGACCCCTGTGACAGCGGAGCCGCG	119
Db	942	AGCCCGCCGACAGGGGACGCGCCCCACAGCGCCAGGGACCCCTGTGACAGCGGAGCCGCG	10010
OY	120	GGTGGAGGTTATGATTCACAGGGGGCGCCCCCGGGGCGGTGCTCCCGCGGCTCTGCCGCGT	179
Db	1002	GGTGGAGGTTATGATTCACAGGGGGCGCCCCCGGGGCGGTGCTCCCGCGGCTCTGCCGCGT	10610
OY	180	GCTGGTGCCTGCTGAACCCGCGCGGCGGCAAGGCAAGGCGCTTGCAGCTTTCGGAGTCA	239
Db	1082	GCTGGTGCCTGCTGAACCCGCGCGGCGGCAAGGCAAGGCGCTTGCAGCTTTCGGAGTCA	11210
OY	240	CGTGACAGCCCTTTTGGCTGAGGCTGAAATCTTCCTCACGCTGATGCTCCTGAGCGCGCG	299
Db	1122	CGTGACAGCCCTTTTGGCTGAGGCTGAAATCTTCCTCACGCTGATGCTCCTGAGCGCGCG	11810
OY	300	GAACCCAGCGCGGGAGCTGGTGGGGTGGAGGAGCTGGGCGCGCTGGAGCGCTCTGGTGGT	359
Db	1182	GAACCCAGCGCGGGAGCTGGTGGGGTGGAGGAGCTGGGCGCGCTGGAGCGCTCTGGTGGT	12410
OY	360	CATGCTGGAAGACGGGCTGATATGACAGGAGTGGAGAACGGGCTCATGGAACGGGCTGAGCTG	419
Db	1242	CATGCTGGAAGACGGGCTGATATGACAGGAGTGGAGAACGGGCTCATGGAACGGGCTGAGCTG	13010
OY	420	GGAGACCGCCCATCCAGAAAGCCCTGTGTAGCTCCGACGAGGCGCTGGACACGCGCTGGC	479
Db	1302	GGAGACCGCCCATCCAGAAAGCCCTGTGTAGCTCCGACGAGGCGCTGGACACGCGCTGGC	13610
OY	480	AGCTTCCTTTGAACCATTTATGCTGGCTATGAGCAGGTACCAATGAAGACCTCTCTGACCA	539
Db	1362	AGCTTCCTTTGAACCATTTATGCTGGCTATGAGCAGGTACCAATGAAGACCTCTCTGACCA	14210
OY	540	CTGCACGCTATTGTGTGGCGCGCGGGGTGATACCCATGAAGACTGCTGTCTGACAC	599
Db	1422	CTGCACGCTATTGTGTGGCGCGGGGTGATACCCATGAAGACTGCTGTCTGCTGACAC	14810
OY	600	GACCTTCGGGGCTGGCGCTCTTCTGTGTGCTCAGACTGGGCTGGGGCTCATTTGCTGATGT	659
Db	1482	GACCTTCGGGGCTGGCGCTCTTCTGTGTGCTCAGACTGGGCTGGGGCTCATTTGCTGATGT	15410
OY	660	GGACCTTAGAGGTAGAGATATGGGGCTCTGGGGAGATGCGCTTCACCTCTGGGACCTT	719
Db	1542	GGACCTTAGAGGTAGAGATATCGCGCTCTGGGGAGATGCGCTTCACCTCTGGGACCTT	16010
OY	720	CCTGGCTTGGCAGCCCTGGGACCTTACGCGGGCGCACTGAGGCTTACTCTCCGTATGGAG	779
Db	1602	CCTGGCTTGGCAGCCCTGGGACCTTACGCGGGCGCACTGAGGCTTACTCTCCGTATGGAG	16610


```

Db 1962 CTACTTGATATATGTCGCCGTCGTCGCCCTTCGCTTGAGCCCAAGATGGAGAGCTGT 2021
OY 1140 GTTTGCGATGATGGGGAATTTGATGTTTACGAGGCCGTGACAGGCCAGTGCACCCAAA 1199
Db 2022 GTTTGCGATGATGGGGAATTTGATGTTTACGAGGCCGTGACAGGCCAGTGCACCCAAA 2081
OY 1200 CTACTTCTGATGATGTCAGCGGTTGCTGAGACCCGCCAGCTGAGAGCCGCCAGCAGAT 1259
Db 2082 CTACTTCTGATGATGTCAGCGGTTGCTGAGACCCGCCAGCTGAGAGCCGCCAGCAGAT 2141
OY 1260 GCCACGCCAGAGAGCCCTTATGACCCCTGAGGCCGCTGTGCTTATGATGTCCTACTTGC 1319
Db 2142 GCCACGCCAGAGAGCCCTTATGACCCCTGAGGCCGCTGTGCTTATGATGTCCTACTTGC 2201
OY 1320 AGGACCCCTTCTCTCCCTAGGCGTCAGGCGCCCTGTCCACAGCTCTGTGAGGCTGAG 1379
Db 2202 AGGACCCCTTCTCTCCCTAGGCGTCAGGCGCCCTGTCCACAGCTCTGTGAGGCTGAG 2261
OY 1380 GAGACTCTCTGAGAGAGGTCAGAGGTCAGAGCTATGCTTGGGAGGACAGGCGAGAA 1439
Db 2262 GAGACTCTCTGAGAGAGGTCAGAGGTCAGAGCTATGCTTGGGAGGACAGGCGAGAA 2321
OY 1440 TGAAGTCTCTGAGAGCCCGACGCTGAGGCCCACTGCTATGTAAGGCTTCTAG 1499
Db 2322 TGAAGTCTCTGAGAGCCCGACGCTGAGGCCCACTGCTATGTAAGGCTTCTAG 2381
OY 1500 TTTGTTCTGAGACCCCGACGCTGAGGCCCACTGCTATGTAAGGCTTCTAG 1559
Db 2382 TTTGTTCTGAGACCCCGACGCTGAGGCCCACTGCTATGTAAGGCTTCTAG 2441
OY 1560 AAAAAAAAAA 1571
Db 2442 CACAAACATGAA 2453

```

Search completed: May 17, 2003, 15:40:37
Job time : 3804 secs


```
Db 447 CCGTACACCCCTTTTGGTGAAGCTGAATATCCCTTACCTATCTCCTGACGGCGC 506
Oy 300 GAACACGCGCGGAGCTGTCGGTTCGAGAGAGCTGGGCGCTGGAGCGCTGTGGT 359
Db 507 GAACACGCGCGGAGCTGTCGGTTCGAGAGAGCTGGGCGCTGGAGCGCTGTGGT 566
Oy 360 CATGCTGAGAGAGCGGCTGATCAGAGAGTGTGAACGGGCTCATGAGAGCGGCTGAGT 419
Db 567 CATGCTGAGAGAGCGGCTGATCAGAGAGTGTGAACGGGCTCATGAGAGCGGCTGAGT 626
Oy 420 GGAACGCGCGCTCAGAAAGCCCTGTGTAGCTCCAGAGAGCTGTGGAGCGGCTGAGC 479
Db 627 GGAACGCGCGCTCAGAAAGCCCTGTGTAGCTCCAGAGAGCTGTGGAGCGGCTGAGC 686
Oy 480 AGCTTCCTGAGACATTTATGCTGGCTATGAGAGAGTCAACCAATGAAGAGCTTCAGCAA 539
Db 687 AGCTTCCTGAGACATTTATGCTGGCTATGAGAGAGTCAACCAATGAAGAGCTTCAGCAA 746
Oy 540 CTGACGCTATGCTGCGCGCGGCTGCTGACACCATTAACCTGCTGCTGACAC 599
Db 747 CTGACGCTATGCTGCGCGCGGCTGCTGACACCATTAACCTGCTGCTGACAC 806
Oy 600 GCGTTCGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 807 GCGTTCGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Oy 660 GGACCTAGAGAGTGAAGTATCGGCGCTGCGGAGAGATCGCTTCACTCTGGGACCTT 719
Db 867 GGACCTAGAGAGTGAAGTATCGGCGCTGCGGAGAGATCGCTTCACTCTGGGACCTT 926
Oy 720 CCTGCTGCTGCGAGCGCTGCGACCTTACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
Db 927 CCTGCTGCTGCGAGCGCTGCGACCTTACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
Oy 780 AGTGGTTCACAAACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
Db 987 AGTGGTTCACAAACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
Oy 840 CCTGCTGCGACCTGCGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899
Db 1047 CCTGCTGCGACCTGCGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
Oy 900 TGTGCTGCTGCTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
Db 1107 TGTGCTGCTGCTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1166
Oy 960 GGGCGCTGCGAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
Db 1167 GGGCGCTGCGAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1226
Oy 1020 CATGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
Db 1227 CATGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1286
Oy 1080 CTACTTGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1139
Db 1287 CTACTTGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1346
Oy 1140 GTTTCAGTGTATGAGGAGTGTATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1199
Db 1347 GTTTCAGTGTATGAGGAGTGTATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1406
Oy 1200 CTACTTGTATGAGGAGTGTATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259
Db 1407 CTACTTGTATGAGGAGTGTATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1466
Oy 1260 GCGACGCGCGAGAGAGCTTATGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1319
Db 1467 GCGACGCGCGAGAGAGCTTATGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1526
Oy 1320 AGGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1379
```

```
Db 1527 AGGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1586
Oy 1380 GAGACTCTCTCTGAGAGAGGCTGAGAGAGTGTAGCTCTTGGGGGAGAGCGCACAA 1439
Db 1587 GAGACTCTCTCTGAGAGAGGCTGAGAGAGTGTAGCTCTTGGGGGAGAGCGCACAA 1446
Oy 1440 TGAAGTCTGAGGCTGAGAGAGGCTGAGAGAGTGTAGCTCTTGAAGGCTTCTAG 1499
Db 1647 TGAAGTCTGAGGCTGAGAGAGGCTGAGAGAGTGTAGCTCTTGAAGGCTTCTAG 1706
Oy 1500 TTTGTTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1559
Db 1707 TTTGTTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1766
Oy 1560 AAAAAAAAAAAAAA 1573
Db 1767 AAAAAAAAAAAAAA 1780

RESULT 2
PCT-US03-02588-25
Sequence 25, Application PC/TUS0302588
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Hunter, John Joseph
APPLICANT: Macbeth, Kyle J.
APPLICANT: Tsai, Fong-Ying
APPLICANT: Lesoon, Andrea
APPLICANT: Lightcap, Eric S.
APPLICANT: Williamson, Mark
APPLICANT: Rudolph-Owen, Laura A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1884, 2086, 32427, 2160,
TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
FILE REFERENCE: MP102-020
CURRENT FILING DATE: 2003-01-30
PCT-US03-02588
PRIOR FILING DATE: 2002-01-31
PRIOR FILING DATE: 2002-01-31
PRIOR FILING DATE: 2002-03-15
PRIOR FILING DATE: 2002-03-15
PRIOR FILING DATE: 2002-04-09
PRIOR FILING DATE: 2002-04-09
PRIOR FILING DATE: 2002-04-10
PRIOR FILING DATE: 2002-04-16
PRIOR FILING DATE: 2002-04-16
PRIOR FILING DATE: 2002-04-19
PRIOR FILING DATE: 2002-05-24
PRIOR FILING DATE: 2002-05-24
PRIOR FILING DATE: 2002-05-31
PRIOR FILING DATE: 2002-06-14
PRIOR FILING DATE: 2002-06-17
Remaining Prior Application data removed - see file wrapper or PALM.
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 25
LENGTH: 1799
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (359)...(1513)
PCT-US03-02588-25
```

Query Match 98.5%; Score 1549.8; DB 1; Length 1799;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1562; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 1 GCGCCACACCGCCCTTGGAGAGCCCGCTGGGAGACACCGATAGAGACTGAGAGCAG 60
 Db 229 GCGCCACACCGCCCTTGGAGAGCCCGCTGGGAGACACCGATAGAGACTGAGAGCAG 288
 QY 61 ACCCGCCGCGAGGAGAGG-CCCCACAGCGCAGAGACCCCTGGACAGCGAGCGG 119
 Db 289 ACCCGCCGCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 348
 QY 120 GGTGAGGTTATGATGACAGCGGCGGCGCCCGGCGGCTGCTCCGCGGCGCTGCCGCT 179
 Db 349 GGTGAGGTTATGATGACAGCGGCGGCGCCCGGCGGCTGCTCCGCGGCGCTGCCGCT 408
 QY 180 GGTGAGGTTATGATGACAGCGGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 239
 Db 409 GGTGAGGTTATGATGACAGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 468
 QY 240 GGTGAGGTTATGATGACAGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 299
 Db 469 GGTGAGGTTATGATGACAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 528
 QY 300 GAACACGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 359
 Db 529 GAACACGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 588
 QY 360 CATGTCGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 419
 Db 589 CATGTCGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 648
 QY 420 GGAGACCGCGCATGCAAGAGCCCGCTGTGTAGAGCCCGAGAGGAGGAGGAGGAGGAGG 479
 Db 649 GGAGACCGCGCATGCAAGAGCCCGCTGTGTAGAGCCCGAGAGGAGGAGGAGGAGGAGG 708
 QY 480 AGCTTCCTTGAACCAATATGCTGGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 539
 Db 709 AGCTTCCTTGAACCAATATGCTGGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 768
 QY 540 CTGACGAGCTATTGCTGCTGCGCGCGGCTGCTGTCACCATGAACTCTGCTGACAC 599
 Db 769 CTGACGAGCTATTGCTGCTGCGCGCGGCTGCTGTCACCATGAACTCTGCTGACAC 828
 QY 600 GGGTGGGGGCGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 659
 Db 829 GGGTGGGGGCGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 888
 QY 660 GGACCTAGAGAGTGAAGATATGCGGCTGTGGGAGAGATGCGCTTACCTGTGGACACT 719
 Db 889 GGACCTAGAGAGTGAAGATATGCGGCTGTGGGAGAGATGCGCTTACCTGTGGACACT 948
 QY 720 CCTGCTGTGCGAGCGCTTGTGCGACCTACCGGCGGAGTGTGCTTCTGTAGAGAG 779
 Db 949 CCTGCTGTGCGAGCGCTTGTGCGACCTACCGGCGGAGTGTGCTTCTGTAGAGAG 1008
 QY 780 AGTGGTTTCAAGACACTGCGCTCCCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 839
 Db 1009 AGTGGTTTCAAGACACTGCGCTCCCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1068
 QY 840 CCTTGTGCACTGAGAGAGCAGTGTGCTTGTGAGAGAGTGTGCTTGTGAGAGAGT 899
 Db 1069 CCTTGTGCACTGAGAGAGCAGTGTGCTTGTGAGAGAGTGTGCTTGTGAGAGAGT 1128
 QY 900 TGTGCTAGTGTGAGAGTGTGCTGCTGCACTGCGGAGTGTGCTGCTGCACTGCT 959
 Db 1129 TGTGCTAGTGTGAGAGTGTGCTGCTGCACTGCGGAGTGTGCTGCTGCACTGCT 1188
 QY 960 GGGCGGCTGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
 Db 1189 GGGCGGCTGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248

QY 1020 CATGCTGTGCGCGCTTCTGCTGCGCATGAGAGAGGCGAGGAGATGAGATGATGCC 1079
 Db 1249 CATGCTGTGCGCGCTTCTGCTGCGCATGAGAGAGGCGAGGAGATGAGATGATGCC 1308
 QY 1080 CATCTTGTATATGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1139
 Db 1309 CATCTTGTATATGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1368
 QY 1140 GTTTCAGTGTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
 Db 1369 GTTTCAGTGTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1428
 QY 1200 CTACTTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
 Db 1429 CTACTTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1488
 QY 1260 GCGACCGCGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1319
 Db 1489 GCGACCGCGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1548
 QY 1320 AGGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1379
 Db 1549 AGGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1608
 QY 1380 GAGACTCTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1439
 Db 1609 GAGACTCTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1668
 QY 1440 TGAAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1499
 Db 1669 TGAAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1728
 QY 1500 TTTGTTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1559
 Db 1729 TTTGTTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1788
 QY 1560 AAAAAAAAAA 1570
 Db 1789 AAAAAAAAAA 1799

RESULT 3
 US-10-354-25
 ; Sequence 25, Application US/10354358
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc
 ; APPLICANT: Hunter, John Joseph
 ; APPLICANT: Macbeth, Kyle J.
 ; APPLICANT: Tsai, Fong-Ying
 ; APPLICANT: Lesoon, Andrea
 ; APPLICANT: Lightcap, Eric S.
 ; APPLICANT: Williamson, Mark
 ; APPLICANT: Rudolph-Owen, Laura A.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
 ; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
 ; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 16842, 17667, 9235,
 ; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
 ; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
 ; TITLE OF INVENTION: 8890, 2100, 9288, 64698, 10460, 20893, 33230, 1586, 9943,
 ; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
 ; FILE REFERENCE: MP102-020P.RNOMNIM
 ; CURRENT APPLICATION NUMBER: US/10354,358
 ; CURRENT FILING DATE: 2003-01-30
 ; PRIOR APPLICATION NUMBER: US 60/353,600
 ; PRIOR FILING DATE: 2002-01-31
 ; PRIOR APPLICATION NUMBER: US 60/364,517
 ; PRIOR FILING DATE: 2002-03-15
 ; PRIOR APPLICATION NUMBER: US 60/371,075
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: US 60/371,507
 ; PRIOR FILING DATE: 2002-04-10
 ; PRIOR APPLICATION NUMBER: US 60/372,984

; PRIOR FILING DATE: 2002-04-16
 ; PRIOR APPLICATION NUMBER: US 60/374,194
 ; PRIOR FILING DATE: 2002-04-19
 ; PRIOR APPLICATION NUMBER: US 60/382,995
 ; PRIOR FILING DATE: 2002-05-24
 ; PRIOR APPLICATION NUMBER: US 60/385,023
 ; PRIOR FILING DATE: 2002-05-31
 ; PRIOR APPLICATION NUMBER: US 60/388,853
 ; PRIOR FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: US 60/389,395
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 122
 ; SOFTWARE: FASTSeq for Windows Version 4.0
 ; SEQ ID NO 25
 ; LENGTH: 1799
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (359)...(1513)
 US-10-354-358-25

Query Match 98.5%; Score 1549.8; DB 9; Length 1799;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1562; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

Oy 1 GCGCCACAGCGCGCGCTGGAGCGCGCGCTGGAGCGCGCGATAGAGCGCGCGAGG 60
 Db 229 GCGCCACAGCGCGCGCTGGAGCGCGCGCTGGAGCGCGCGATAGAGCGCGAGG 288
 Oy 61 AGCGCGCGCGCGCGCGAGCG -CGCCAGCGCGCGCGCGCGCGCGCGCGCGCG 119
 Db 289 AGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 348
 Oy 120 GGTGAGCG 179
 Db 349 GGTGAGCG 408
 Oy 180 GGTGAGCG 239
 Db 409 GGTGAGCG 468
 Oy 240 GGTGAGCG 299
 Db 469 GGTGAGCG 528
 Oy 300 GGTGAGCG 359
 Db 529 GGTGAGCG 588
 Oy 360 GGTGAGCG 419
 Db 589 GGTGAGCG 648
 Oy 420 GGTGAGCG 479
 Db 649 GGTGAGCG 708
 Oy 480 GGTGAGCG 539
 Db 709 GGTGAGCG 768
 Oy 540 GGTGAGCG 599
 Db 769 GGTGAGCG 828
 Oy 600 GGTGAGCG 659
 Db 829 GGTGAGCG 888
 Oy 660 GGTGAGCG 719

Db 889 GGTGAGCG 948
 Oy 720 GGTGAGCG 779
 Db 949 GGTGAGCG 1008
 Oy 780 GGTGAGCG 839
 Db 1009 GGTGAGCG 1068
 Oy 840 GGTGAGCG 899
 Db 1069 GGTGAGCG 1128
 Oy 900 GGTGAGCG 959
 Db 1129 GGTGAGCG 1188
 Oy 960 GGTGAGCG 1019
 Db 1189 GGTGAGCG 1248
 Oy 1020 GGTGAGCG 1079
 Db 1249 GGTGAGCG 1308
 Oy 1080 GGTGAGCG 1139
 Db 1309 GGTGAGCG 1368
 Oy 1140 GGTGAGCG 1199
 Db 1369 GGTGAGCG 1428
 Oy 1200 GGTGAGCG 1259
 Db 1429 GGTGAGCG 1488
 Oy 1260 GGTGAGCG 1319
 Db 1489 GGTGAGCG 1548
 Oy 1320 GGTGAGCG 1379
 Db 1549 GGTGAGCG 1608
 Oy 1380 GGTGAGCG 1439
 Db 1609 GGTGAGCG 1668
 Oy 1440 GGTGAGCG 1499
 Db 1669 GGTGAGCG 1728
 Oy 1500 GGTGAGCG 1559
 Db 1729 GGTGAGCG 1788
 Oy 1560 GGTGAGCG 1570
 Db 1789 GGTGAGCG 1799

RESULT 4
 US-09-949-016-3940
 ; Sequence 3940, Application US/09949016
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. CRAIG ET AL.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: C1001307
 ; CURRENT FILING DATE: 2000-04-14

```

: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FASTSEQ for Windows Version 4.0.
: SEQ ID NO 3940
: LENGTH: 1764
: TYPE: DNA
: ORGANISM: Human
: OS-09-949-016-3940

```

Query Match	92.8%;	Score 1459.8;	DB 7;	Length 1764;
Best Local Similarity	97.3%;	Pred. No. 8e-312;		
Matches 1517; Conservative	0;	Mismatches 37;	Indels 5;	Gaps 3

QY	1	GCCCACAGCGCGGCGCTGGCAGCCCGCGCTGGGACACACGATTAAGAGAGCTAAAGCAGG	60
Db	207	GCCCCAACAGCGGCGCTGGCAGCCCGCGCTGGGACACACGATTAAGAGAGCTAAAGCAGG	266
QY	61	AGCCGCGCCACAGGGGAGCG - CCGCACAGCGCCAGGGAGCCCTCTGACAGCGGGAGCCGCG	119
Db	267	AGCCGCGCCACAGGGGAGCGCCCGCCACAGCGCCAGGGAGCCCTCTGACAGCGGGAGCCGCG	326
QY	120	GGTGAAGATTATGGATTCACAGGGGCGGCCCCCGGCGGCTGTCTCCGCGGCGCTCGCCGGT	179
Db	327	GGTGAAGATTATGGATTCACAGAGGGGCGGCCCCCGGCGGTAGCTCCGCGGCGCTTTCCTA	386
QY	180	GCTGTGCTCTCTGAAACCCGCGGGGCGGACAGGAGCGCTTGACAGCTTTCGGAGTCA	239
Db	387	TGAATGACTGTGAAAC - GCGAAGGCAAGGCAAGG - TTCAAGCTTTCGGAGTCA	442
QY	240	CGTGCAGCCCCCTTTTGGCTGAGGCTGAAATCTCTTACGCTGAATGCTACAGTGAAGCGCG	299
Db	443	CGTGCAGCCCCCTTTTGGCTGAGGCTGAAATCTCTTACGCTGAATGCTACAGTGAAGCGCGCG	502
QY	300	GAACACAGCGCGGGAGCTGTGTGGCGGCGAGGAGCTGGGCGCGCTGGAGCGCTGTGTGT	359
Db	503	GAACACAGCGCGGAGACTGTGTGGATGGGAAAGCTGGGCGCGTGAAGAGCTGTGTGTGT	562
QY	360	CATGTCTGAGAGACGGGCTGTATGACAGAGGTGTGAACGGGCTCATAGAGCGGCTGAGCTG	419
Db	563	CATGTCTGAGAGAGCGGCTGTATGACAGAGGTGTGAACGGGCTCATAGAGCGGCTGAGCTG	622
QY	420	GGAACACGGCCATTCAGAAAGCCCTGTGTAGCTTCCAGCAGGCTCTGGACAGCGCTGGC	479
Db	623	GGAACACGGCCATTCAGAAAGCCCTGTGTAGCTTCCAGCAGGCTCTGGGACAGCGCTGGC	682
QY	480	AGCTTTCCTTGAACCATTAATGCTGGATATGAGCAGGAGTACCATGAAAGATCTCTGACCA	539
Db	683	AGCTTTCCTTGAACCATTAATGCTGGATATGAGCAGGAGTACCATGAAAGATCTCTGACCA	742
QY	540	CTGACAGCTATTGCTGTGCCCGCGGCTGTGTACCCATGAACCTGTGTCTCTGACAC	599
Db	743	CTGACAGCTATTGCTGTGCCCGCGGCTGTGTACCCATGAACCTGTGTCTCTGACAC	802
QY	600	GGCTTTCGGGGGCTGGGCGCTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	659
Db	803	GGCTTTCGGGGGCTGGGCGCTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	862
QY	660	GGACCTTAGAGAGTGAAGAATATGGGGGCTGGGGGAGATAGCGTTACATCTGGGGACCTT	719
Db	863	GGACCTTAGAGAGTGAAGAATATGGGGGCTGGGGGAGATAGCGTTACATCTGTGGACACTT	922
QY	720	CTGTGCTGTGGCAGCCCTGTGCGACCTAACCGCGGCGAGTGGCTTACCTCTGTAGAGAG	779
Db	923	CTGTGCTGTGGCAGCCCTGTGCGACCTAACCGCGGCGAGTGGCTTACCTCTGTAGAGAG	982
QY	780	AGTGGGTTCCAAAGACACTCTCTCCCGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	839
Db	983	AGTGGGTTCCAAAGACACTCTCTCCCGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1042

```

840 CCGTGTGCACTGGAGGAGGCCAGTGGCCCTTCATCTGACATGAGTGTGGCCGACAGGAGCACTT 899
Db 1043 CTTTGTGCACTGGAGGAGGCCAGTGGCCCTTCATCTGACATGAGTGTGGCCGACAGGAGCACTT 110
QY 900 TGTGCTACTCTGTGGACATGCTGTCACTGTGCACCTTGGGCAATGAGATGTTTGTCTGCACCAT 959
Db 1103 TGTGCTACTCTGTGGACATGCTGTGCACCTGTGCACCTTGGGCAATGAGATGTTTGTCTGCACCAT 116
QY 960 GGGCCGCTGTGCAGCTGGCGCTCATGATCTGTGTTCTACGTGCGGGCGGGAGTGTCTGTGCG 101
Db 1163 GGGCGCGCTGTGCAGCTGGCGCTCATGATCTGTGTTCTACGTGCGGGCGGGAGTGTCTGTGCG 122
QY 1020 CATGCTGTGGGCGCCTCTTCCGTGGGCATGGAGAAAGGCGAGCATATGAGAGTATGATAGGCC 107
Db 1223 CATGCTGTGGGCGCCTCTTCCGTGGGCATGGAGAAAGGCGAGCATATGAGATATGATAGGCC 128
QY 1080 CTACTTGGTATATGTCGCCCTGTGTCGCGCTTCCGCTTGGAGCCCAAGATGGGAAAGGTGT 113
Db 1283 CTACTTGGTATATGTCGCCCTGTGTCGCGCTTCCGCTTGGAGCCCAAGATGGGAAAGGTGT 134
QY 1140 GTTTTCAGTGGATGGGGAATTGATGCTTACGAGAGCCGTGCGAGGCGCAGGTGCACCCAAA 119
Db 1343 GTTTTCAGTGGATGGGGAATTGATGCTTACGAGAGCCGTGCGAGGCGCAGGTGCACCCAAA 140
QY 1200 CTACTTTCGAGTATGGTCACAGGCTGTGCGTGGAGACCCCGCCCAAGCTGGAAGCCCCAGCAGAT 125
Db 1403 CTACTTTCGAGTATGGTCACAGGCTGTGCGTGGAGACCCCGCCCAAGCTGGAAGCCCCAGCAGAT 146
QY 1260 GCCACCGCCAGAGAAGCCCTTATGACCCCTGGGCGGCGCTGTGCTTAGTGTCTACTTGC 131
Db 1463 GCCACCGCCAGAGAAGCCCTTATGACCCCTGGGCGGCGCTGTGCTTAGTGTCTACTTGC 152
QY 1320 AGGACCCCTTCCCTCCTTCCCTTATGAGGCTGCAGGGCCCTGTCCACAGCTCCTGTGTGGGGTGGAG 137
Db 1523 AGGACCCCTTCCCTCCTTCCCTTATGAGGCTGCAGGGCCCTGTCCACAGCTCCTGTGTGGGGTGGAG 158
QY 1380 GAGACTCTCTGGAGAAAGGATGAGAAAGCTGAGAGCTATGCTTTTGGGGGACAGGCCAGAA 143
Db 1583 GAGACTCTCTGGAGAAAGGATGAGAAAGCTGAGAGCTATGCTTTTGGGGGACAGGCCAGAA 164
QY 1440 TGAAGTCTCTGGGTCAGAGCCCAAGCTGTGGTGGGCCCAAGCTGTATGTAAGGCCCTTCTAG 149
Db 1643 TGAAGTCTCTGGGTCAGAGCCCAAGCTGTGGTGGGCCCAAGCTGTATGTAAGGCCCTTCTAG 170
QY 1500 TTTGTTCTGAGACCCGCCCAAGCCAGAACCAATTCCAATTAAGTGAATTTCCACGCTG 1558
Db 1703 TTTGTTCTGAGACCCGCCCAAGCCCAAGAACCAATTCCAATTAAGTGAATTTCCACGCTG 1761

RESULT 5
US-10-264-237-1180/c
: Sequence 1180, Application US/10264237
: GENERAL INFORMATION:
: APPLICANT: Birse et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PA13P1
: CURRENT APPLICATION NUMBER: 2002-10-04
: PRIOR APPLICATION NUMBER: PCT/US01/16450
: PRIOR FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: US 60/205,515
: NUMBER OF SEQ ID NOS: 2876
: SOFTWARE: PatentIn Ver. 3.1
: SEQ ID NO 1180
: LENGTH: 1562
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (7)..(77)
: OTHER INFORMATION: n equals a,t,g, or c

```


Db 446 --CGCGAAGGCAAGGCGCAAGC--TTACAGCTCTTCGAGATCACTGTGACGCCCTTTTGG 501
OY 257 CTGAGGCTGAATCTCTCTTACAGCTGATGCTCAGTGAGCGGCGGAACCAAGCGCGGAGG 316
Db 502 CTGAGGCTGAATCTCTCTTACAGCTGATGCTCAGTGAGCGGCGGAACCAAGCGCGGAGG 561
OY 317 TGGTGGGTCGAGAGAGCTGGGCGGCTGGGAGCGCTGGTGGTCACTGTGAGAGCGG 376
Db 562 GGTGGGATCGGAGAGCTGGGCGGCTGGGAGCGCTGGTGGTCACTGTGAGAGCGG 621
OY 377 TGATGACAGAGGTGGTGAACGGGCTGATGAGAGCGGCTGATGAGAGCGGCTGATGAG 436
Db 622 TGATGACAGAGGTGGTGAACGGGCTGATGAGAGCGGCTGATGAGAGCGGCTGATGAG 661
OY 437 AGCCCTGTGTAGAGCTCCAGAGAGGCTGAGAGCGGCTGAGAGCGGCTGATGAGAGCT 496
Db 682 AGCCCTGTGTAGAGCTCCAGAGAGGCTGAGAGCGGCTGAGAGCGGCTGATGAGAGCT 741
OY 497 ATGCTGGTATGAGAGAGGTCACCAATGAAGACCTCTGACCACTGACAGCTATTCCTGT 556
Db 742 ATGCTGGTATGAGAGAGGTCACCAATGAAGACCTCTGACCACTGACAGCTATTCCTGT 801
OY 557 GCGCGCGGCTGCTGTACACCATGAACCTGCTGTCTGACAGAGCGCTGGGCGCTGCGC 616
Db 802 GCGCGCGGCTGCTGTACACCATGAACCTGCTGTCTGACAGAGCGCTGGGCGCTGCGC 861
OY 617 TCTTCTGTGTGTGCTGAGCTGGGCGCTGAGTGGGCTGATGATGATGATGATGATGATG 676
Db 862 TCTTCTGTGTGTGCTGAGCTGGGCGCTGAGTGGGCTGATGATGATGATGATGATGATG 921
OY 677 AGTATGCGGCTGTGGGAGAGTGGGCTGCTGCTGAGGACCTTCTGCTGCTGAGGAGG 736
Db 922 AGTATGCGGCTGTGGGAGAGTGGGCTGCTGCTGAGGACCTTCTGCTGCTGAGGAGG 981
OY 737 TGGGACACACCGCGGCGGCTGAGTGGGCTGCTGCTGAGGAGAGGCTGCTGAGGAGG 796
Db 982 TGGGACACACCGCGGCGGCTGAGTGGGCTGCTGCTGAGGAGAGGCTGCTGAGGAGG 1041
OY 797 CTGCTCCCGGCTGTGTGTGCTGAGAGGCGGCTGATGAGGACCTTCTGCTGCTGAGG 856
Db 1042 CTGCTCCCGGCTGTGTGTGCTGAGAGGCGGCTGATGAGGACCTTCTGCTGCTGAGG 1101
OY 857 AGCAGCTGCGCTCTCACTGAGAGTGGTGGCGGAGAGGCTTGTGTGTGTGTGTGTGTGT 916
Db 1102 AGCAGCTGCGCTCTCACTGAGAGTGGTGGCGGAGAGGCTTGTGTGTGTGTGTGTGTGT 1161
OY 917 TGTGTGACACGACCTGGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 976
Db 1162 TGTGTGACACGACCTGGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1221
OY 977 GCGTCATGATCTGTCTGAGTGGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1036
Db 1222 GCGTCATGATCTGTCTGAGTGGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1281
OY 1037 TCTGTGACATGAGAGAGGAGCATATGAGATGATGATGATGATGATGATGATGATGAT 1096
Db 1282 TCTGTGACATGAGAGAGGAGCATATGAGATGATGATGATGATGATGATGATGATGAT 1341
OY 1097 CCGTGTGCTCTGCGCTGT 1156
Db 1342 CCGTGTGCTCTGCGCTGT 1401
OY 1157 AATTGATGATGAGAGAGGAGCATATGAGATGATGATGATGATGATGATGATGATGAT 1216
Db 1402 AATTGATGATGAGAGAGGAGCATATGAGATGATGATGATGATGATGATGATGATGAT 1461
OY 1217 GCGGTTGCTGTGAGAGCGGCGGCTGAGAGCGGCTGAGAGCGGCTGAGAGCGGCTGAG 1276
Db 1462 GCGGTTGCTGTGAGAGCGGCGGCTGAGAGCGGCTGAGAGCGGCTGAGAGCGGCTGAG 1521
OY 1277 CATTATGAGAGCGGCGGCTGAGAGCGGCTGAGAGCGGCTGAGAGCGGCTGAGAGCGG 1336
Db 1522 CATTATGAGAGCGGCGGCTGAGAGCGGCTGAGAGCGGCTGAGAGCGGCTGAGAGCGG 1581

OY 1337 CTTAGGAGCTGAGAGGCTGTGACAGCTCTGAGGAGTGGAGAGAGCTGCTGTGAGAA 1396
Db 1582 CTTAGGAGCTGAGAGGCTGTGACAGCTCTGAGGAGTGGAGAGAGCTGCTGTGAGAA 1641
OY 1397 GGTGTGAGAGGCTGAGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1456
Db 1642 GGTGTGAGAGGCTGAGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1701
OY 1457 AGCCAGCTGCTGGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1516
Db 1702 AGCCAGCTGCTGGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1761
OY 1517 ACCCCAGAGAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1558
Db 1762 ACCCCAGAGAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1803

RESULT 7

US-60-453-135-5904

Sequence 5904, Application US/60453135

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: IAKOBYOVA, Olga

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

CURRENT APPLICATION NUMBER: US/60/453, 135

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5904

LENGTH: 1806

TYPE: DNA

ORGANISM: Homo sapiens

US-60-453-135-5904

Query Match

Best Local Similarity 89.8%; Score 1412; DB 11; Length 1806;

Matches 1523; Conservative 5; Mismatches 25; Indels 49; Gaps 5;

OY 1 GCCCCACAGCGCGCCCTGCGAGCGCCGCTGGGAGACGAGATGAAGAGCTGAGGAGG 60
Db 207 GCCCCACAGCGCGCCCTGCGAGCGCCGCTGGGAGACGAGATGAAGAGCTGAGGAGG 266
OY 61 AGCCGCGCGCCAGCGGAGAGG-CCCCACAGCGCCAGGAGACCCCTGGAGCGGAGCGCG 119
Db 267 AGCCGCGCGCCAGCGGAGAGG-CCCCACAGCGCCAGGAGACCCCTGGAGCGGAGCGCG 326
OY 120 GGTGAGAGTTATGATCCAG-----139
Db 327 GGTGAGAGTTATGATCCAGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 386
OY 140 ---CGGCGGCGCCCGGCGGCTGCTCCGCGGCGCTGCGCGGCTGCTGCTGCTGAGC 196
Db 387 AGCAGGCGGCGCCCGGCGGCTGAGTCCCGGCGGCTGCTGCTGCTGAGC-AGTACTGCTGAGC 445
OY 197 CCGGCGGCGGAGGAGGAGGCTTGTGAGTCTTTCGGAATGACAGTGTGAGTGTGAGTGT 256
Db 446 --CGCGAAGGCAAGGCGCAAGG---TTCAAGCTTTCGGAATGACAGTGTGAGTGTGAGTGT 501
OY 257 CTGAGGCTGAATCTCTTACAGCTGATGCTCACTGAGCGGCGGAACCAAGCGGCGGAGG 316
Db 502 CTGAGGCTGAATCTCTTACAGCTGATGCTCACTGAGCGGCGGAACCAAGCGGCGGAGG 561
OY 317 TGGTGGGTCGAGAGAGCTGGGCGGCTGGGAGCGCTGGTGGTCACTGTGAGAGCGG 376
Db 562 GGTGGGATCGGAGAGCTGGGCGGCTGGGAGCGCTGGTGGTCACTGTGAGAGCGG 621
OY 377 TGATGACAGAGGTGGTGAACGGGCTGATGAGAGCGGCTGATGAGAGCGGCTGATGAG 436
Db 622 TGATGACAGAGGTGGTGAACGGGCTGATGAGAGCGGCTGATGAGAGCGGCTGATGAG 681


```
Db 862 TCTCTCTGCTCAGCCCTGGCGCTTCATTCGTGATGTGACCTAGAGAGTGA 921
Oy 677 AGATGCGGCTCTGGGGAGATGCGCTTCACTCTGGGACCTTCCGCGTGGAGGCC 736
Db 922 AGTATGCGCTCTGGGGAGATGCGCTTCACTCTGGGACCTTCCGCGTGGAGGCC 981
Oy 737 TCGGACCTACCGCGGCGGCTTACCTCCCTGTAGAGAGTGGGTTCACAGAC 796
Db 982 TCGGACCTACCGCGGCGGCTTACCTCCCTGTAGAGAGTGGGTTCACAGAC 1041
Oy 797 CTGCTCCCGGCTGTGTGTTCAGAGAGGCCCGGTAGATGACACCTTGTGCCAGT 856
Db 1042 CTGCTCCCGGCTGTGTGTTCAGAGAGGCCCGGTAGATGACACCTTGTGCCAGT 1101
Oy 857 AGCGAGTGCCTCTCACTGAGAGTGTGCGGAGAGGACCTTGTGATGTCAGTCCGAC 916
Db 1102 AGCGAGTGCCTCTCACTGAGAGTGTGCGGAGAGGACCTTGTGATGTCAGTCCGAC 1161
Oy 917 TCGTCACTCGGACCTGGGAGTGAATGTTGCTGACCCATGGGCGCGCTGTGACGCTG 976
Db 1162 TCGTCACTCGGACCTGGGAGTGAATGTTGCTGACCCATGGGCGCGCTGTGACGCTG 1221
Oy 977 GCGTCACTGATCTGTTCTACGTCGCGGCGGAGTGTCTGTGACGTCGTCGCTCT 1036
Db 1222 GCGTCACTGATCTGTTCTACGTCGCGGCGGAGTGTCTGTGACGTCGTCGCTCT 1281
Oy 1037 TCGTCACTGAGAGAGGAGGAGGATGAGTGAATGAGCCCTGATGATGTCG 1096
Db 1282 TCGTCACTGAGAGAGGAGGAGGATGAGTGAATGAGCCCTGATGATGTCG 1341
Oy 1097 CCGTGTGCGCTTCCGCTTGGAGGCCAGAGTGGAAAGTGTGTTGCACTGATGAGG 1156
Db 1342 CCGTGTGCGCTTCCGCTTGGAGGCCAGAGTGGAAAGTGTGTTGCACTGATGAGG 1401
Oy 1157 AATTGATGTTAGCGAGGCGGTGACAGGCGGAGTGCACCCAACTACTTGTGATGTC 1216
Db 1402 AATTGATGTTAGCGAGGCGGTGACAGGCGGAGTGCACCCAACTACTTGTGATGTC 1461
Oy 1217 GCGGTTGCGTGGAGGCCCGGCGGAGTGAAGGCCAGAGTGCACCGCGCGAGAGG 1276
Db 1462 GCGGTTGCGTGGAGGCCCGGCGGAGTGAAGGCCAGAGTGCACCGCGCGAGAGG 1521
Oy 1277 CCTTATGACCCCTGGGCGGCTGTGCTTACTTACTTGGAGAGCCCTTCTCTTTC 1336
Db 1522 CCTTATGACCCCTGGGCGGCTGTGCTTACTTACTTGGAGAGCCCTTCTCTTTC 1581
Oy 1337 CCGAGGCTGCGAGGCGCTTGCACAGCTCCTGTGGGGTGGAGAGTCTCTGAGAA 1396
Db 1582 CCGAGGCTGCGAGGCGCTTGCACAGCTCCTGTGGGGTGGAGAGTCTCTGAGAA 1641
Oy 1397 GCGTGAAGGTTGAGGCTATGCTTGGGGGAGAGGCGAGAAATGAATGCTGGTCCAG 1456
Db 1642 GCGTGAAGGTTGAGGCTATGCTTGGGGGAGAGGCGAGAAATGAATGCTGGTCCAG 1701
Oy 1457 AGCCGAGTGGCTGGGCGGAGCTGCTATGTAAGGCTTCTAGTTGTTGAGACCCC 1516
Db 1702 AGCCGAGTGGCTGGGCGGAGCTGCTATGTAAGGCTTCTAGTTGTTGAGACCCC 1761
Oy 1517 ACCCGAGAGCAATGCAATTAAGTGAATCCAGGCTG 1558
Db 1762 ACCCGAGAGCAATGCAATTAAGTGAATCCAGGCTG 1803
```

RESULT 9

```
US-10-348-052-22
; Sequence 22, Application US/10348052
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; APPLICANT: First, Henrik
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION
; FILE REFERENCE: 200116.405
; CURRENT APPLICATION NUMBER: US/10/348,052
```

```
; CURRENT FILING DATE: 2003-01-17
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-348-052-22
```

Query Match 73.1%; Score 1150.4; DB 9; Length 1152;

Best Local Similarity 99.9%; Pred. No. 1,1e-243;

Matches 1151; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Oy 130 ATGATTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 189
Db 1 ATGATTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
Oy 190 CTGAACCGCGCGCGCGCGCAAGGCGCAAGGCGCTTGCAGCTCTTCCGAGTCACTGACGCC 249
Db 61 CTGAACCGCGCGCGCGCGCAAGGCGCAAGGCGCTTGCAGCTCTTCCGAGTCACTGACGCC 120
Oy 250 CTTTGGCTGAGGCTGAATCTCTTCACTGATGCTCTACTAGCGGCGGAGAACCGCG 309
Db 121 CTTTGGCTGAGGCTGAATCTCTTCACTGATGCTCTACTAGCGGCGGAGAACCGCG 180
Oy 310 CCGGAGCTGCTGCGGCTCGGAGAGTGGGCGGCTGGGAGCGCTGGTGGTCACTGTGGA 369
Db 181 CCGGAGCTGCTGCGGCTCGGAGAGTGGGCGGCTGGGAGCGCTGGTGGTCACTGTGGA 240
Oy 370 GACGGCTGATGACAGAGTGTGAACGGGCTCATGAGAGGCGCTGACCTGAGAGACCGCC 429
Db 241 GACGGCTGATGACAGAGTGTGAACGGGCTCATGAGAGGCGCTGACCTGAGAGACCGCC 300
Oy 430 ATCCAGAGCGGCTGTGTACCTTCCAGCAGGCTGTGGCAACCGCTGGAGCTTCTTG 489
Db 301 ATCCAGAGCGGCTGTGTACCTTCCAGCAGGCTGTGGCAACCGCTGGAGCTTCTTG 360
Oy 490 AACCATTTATGCTGGCTATGAGAGGTCACCAATGAAGACCTCTTGACCACTCAGGCTA 549
Db 361 AACCATTTATGCTGGCTATGAGAGGTCACCAATGAAGACCTCTTGACCACTCAGGCTA 420
Oy 550 TTGCTGTGCGCGCGGCTGTGTCACCCATGAACCTGTCTCTGTGACAGCGCTTGGGG 609
Db 421 TTGCTGTGCGCGCGGCTGTGTCACCCATGAACCTGTCTCTGTGACAGCGCTTGGGG 480
Oy 610 CTGCGCTCTTCTGTGCTGCTACGCTGGCTGGGCTTATTGCTGATGTGACCTAGAG 669
Db 481 CTGCGCTCTTCTGTGCTGCTACGCTGGCTGGGCTTATTGCTGATGTGACCTAGAG 540
Oy 670 AGTGAAGTATGAGGCTGTGGGGGAGATGCGCTTCACTGTGGGACCTTCCGCGCTG 729
Db 541 AGTGAAGTATGAGGCTGTGGGGGAGATGCGCTTCACTGTGGGACCTTCCGCGCTG 600
Oy 730 GCAAGCTTGGGCGCACTACCGCGGCGCACTGAGCTTACCTCTTGAAGAGTGGGTTCC 789
Db 601 GCAAGCTTGGGCGCACTACCGCGGCGCACTGAGCTTACCTCTTGAAGAGTGGGTTCC 660
Oy 790 AAGACACTGCTCCCGCTTGTGTGTCAGAGGCGCGGCTAGATGACACCTTGTGCCA 849
Db 661 AAGACACTGCTCCCGCTTGTGTGTCAGAGGCGCGGCTAGATGACACCTTGTGCCA 720
Oy 850 CTGAGAGGAGCAGTGGCGCTTCACTGAGACGTGTGGCGGAGAGGACTTGGTAGTGC 909
Db 721 CTGAGAGGAGCAGTGGCGCTTCACTGAGACGTGTGGCGGAGAGGACTTGGTAGTGC 780
Oy 910 CTGAGACTGCTGCACTGCACTGCGGCGAGTGAATGTTGCTGCACTAGGCGCGCTGT 969
Db 781 CTGAGACTGCTGCACTGCACTGCGGCGAGTGAATGTTGCTGCACTAGGCGCGCTGT 840
Oy 970 GCAAGCTGCGCTATGATCTGTTC7ACGTCGGGCGGAGTGTCTGTGCACTGCTGCTG 1029
Db 841 GCAAGCTGCGCTATGATCTGTTC7ACGTCGGGCGGAGTGTCTGTGCACTGCTGCTG 900
```

Qy	1030	CGCCTTCTCTGACCATGGAGAAAGGCGACGACATATGAGATGTGAATGCCCACTTTGGTA	1089
Db	901	CGCCTTCTCTGACCATGGAGAAAGGCGACGACATATGAGATGTGAATGCCCACTTTGGTA	960
Qy	1090	TATGTATCCCGTGAGTGGCCTTCCGCTTGGAGCCCAAGCATGSGAAAGTGATGTTGCAGTG	1149
Db	961	TATGTATCCCGTGAGTGGCCTTCCGCTTGGAGCCCAAGCATGSGAAAGTGATGTTGCAGTG	1020
Qy	1150	GATGGGGAATTGATGTTTACGAGAGCGCTGTGCAAGGCGCCAGGTGCACCCAACTACTTTGG	1209
Db	1021	GATGGGGAATTGATGTTTACGAGAGCGCGGTGCAGGCGCCAGGTGCACCCAACTACTTTGG	1080
Qy	1210	ATGTGTACGCGTTGTGCTGTGAGACCCCGCCGACACTGTGAAGCCCCAGCAGATGCCACGSCA	1269
Db	1081	ATGTGTATGAGTGTCGCTGTGAGACCCCGCCGACACTGTGAAGCCCCAGCAGATGCCACGSCA	1140
Qy	1270	GAAGAGCCCTTA	1281
Db	1141	GAAGAGCCCTTA	1152

RESULT 10
US-09-949-016-12897
; Sequence 12897, Application US/09949016

```

1  APPLICANT: VENTER, J. Craig et al.
2  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
3  TITLE OF INVENTION: WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
4  FILE REFERENCE: CLO01307
5  CURRENT APPLICATION NUMBER: US/09/949,016
6  CURRENT FILING DATE: 2000-04-14
7  PRIOR APPLICATION NUMBER: 60/241,755
8  PRIOR FILING DATE: 2000-10-20
9  PRIOR APPLICATION NUMBER: 60/237,768
10 PRIOR FILING DATE: 2000-10-03
11 PRIOR APPLICATION NUMBER: 60/231,498
12 PRIOR FILING DATE: 2000-09-08
13 NUMBER OF SEQ ID NOS: 207012
14 SOFTWARE: FASTSEQ for Windows Version 4.0
15 SEQ ID NO 12897
16 LENGTH: 7195
17 TYPE: DNA
18 ORGANISM: Human
19 US-09-949-016-12897

```

Query Match	67.1%;	Score 1056;	DB 7;	length 7195;
Best Local Similarity	100.0%;	Pred. No. 1,1e-222;		
Matches 1056;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

Oy	503	GCATATAGAGGATACCAATGAAGACCTCTCTGACCACTGAGAGCTATTGCTGTGGCCGC	562
Db	4140	GCATATAGAGGATACCAATGAAGACCTCTCTGACCACTGAGAGCTATTGCTGTGGCCGC	-4199
Oy	563	GGTGGTGTGACCCATGAGACCTGTGTCTGTGACACAGGGTTGGGGGCTGGCGCTTCT	622
Db	4200	GGTGGTGTGACCCATGAGACCTGTGTCTGTGACACAGGGTTGGGGGCTGGCGCTTCT	4255
Oy	623	CTGTGTGACGCTGGCCGTGGGGGCTGATATGTGATGTGGACCTTGAAGAGTGAAGATATC	682
Db	4260	CTGTGTGACGCTGGCCGTGGGGGCTTATTTCTGATGTGTGACCTTGAAGAGTGAAGATATC	4319
Oy	683	GGCGTCTGGGGGAGATGTGCGTTCTACTCTGTGGACACTTCTCTGGCTGTGGACGCGCTTGCGA	742
Db	4320	GGCGTCTGGGGGAGATGTGCGTTCTACTCTGTGGACACTTCTCTGGCTGTGGACGCGCTTGCGA	4379
Oy	743	CTTACCGGGGGGACGTGGCTACCTCCCTGTGTAGAAGAGTGGTGTCCAGACACCTGGCT	802
Db	4380	CTTACCGGGGGGACGTGGCTACCTCCCTGTGTAGAAGAGTGGTGTCCAGACACCTGGCT	4439
Oy	803	CCCCCGTTGTGTCCAGAGAGGGCCGGTGAATGACACACTTGTGCCACTGAGAGACCGAG	862
Db	4440	CCCCCGTTGTGTCCAGAGAGGGCCGGTGAATGACACACTTGTGCCACTGAGAGACCGAG	-4499

OY	863	TTCTCTCAGTGGACAGTGGTGGTGGCAGACAGACTTTGTGCTATGTCCTGGACATCTGG	922
Db	4500	TGCGCTCTCAGTGGACAGTGGTGGTGGCAGACAGACTTTGTGCTATGTCCTGGACATCTGG	4555
OY	923	ACTCGCACTGGGGAGTGAATGTTTGTGTGACCCCATGSGGCGCTGTGACAGCTGGGAGTCA	982
Db	4560	ACTCGCACTGGGGAGTGAATGTTTGTGTGACCCCATGSGGCGCTGTGACAGCTGGGAGTCA	46119
OY	983	TGATCTGTCTACGTCGGGGGGGAGTGTCTGTGGCAATGCTGCTGCTGGCTCTTCCTGG	1042
Db	4620	TGATCTGTCTACGTCGGGGGGGAGTGTCTGTGGCAATGCTGCTGCTGGCTCTTCCTGG	46797
OY	1043	CCATGACAAAGGAGGAGGACATATGAGATATGAAATGCGCTACTTGGTATATGTCGCGG	11020
Db	4680	CCATGACAAAGGAGGAGGACATATGAGATATGAAATGCGCTACTTGGTATATGTCGCGG	47399
OY	1103	TGCGCTTCCGCTGTGAGACCCCAAGATGTGAAAGCTGTGTTTGGACGTGGATGGGAAATGA	11652
Db	4740	TGCGCTTCCGCTGTGAGACCCCAAGATGTGAAAGCTGTGTTTGGACGTGGATGGGAAATGA	47999
OY	1163	TGGTTAGCGAGCGCTGTGAGGCGGACAGGTGACCCCAACTACTTCTGTGATGTGCAGCGTT	12222
Db	4800	TGGTTAGCGAGCGCTGTGAGGCGGACAGGTGACCCCAACTACTTCTGTGATGTGCAGCGTT	48589
OY	1223	GGCTGAGAGCCCCCGCCAGCTGTGAAGCCCCCAGACAGATGCCACGGCCAGAGAGAGCCCTTAT	12822
Db	4860	GGCTGAGAGCCCCCGCCAGCTGTGAAGCCCCCAGACAGATGCCACGGCCAGAGAGAGCCCTTAT	49191
OY	1283	GACCCCTGGGCGCGCTGTGCTTAACTGTCTCATTTGGACGAGCCCTTCTCTCTTCCCTAGG	13422
Db	4920	GACCCCTGGGCGCGCTGTGCTTAACTGTCTCATTTGGACGAGCCCTTCTCTCTTCCCTAGG	49797
OY	1343	GGTGGAGGAGCCTGTCTCACAGCTCTTGTGGGGGTGGAGGAGACTCTCTGTGAGAAGGTGA	14022
Db	4980	GGTGGAGGAGCCTGTCTCACAGCTCTTGTGGGGGTGGAGGAGACTCTCTGTGAGAAGGTGA	50399
OY	1403	GAGGTGGAGGCTATGCTTTGGGGGAGACAGGCGCAGATGAAGTCTTGGGTGAGAGGCCA	14652
Db	5040	GAGGTGGAGGCTATGCTTTGGGGGAGACAGGCGCAGATGAAGTCTTGGGTGAGAGGCCA	50999
OY	1463	GCTGGCTGGGCCAGGCTGCTTATGTAAAGGCTTCTTAAGTTTGTGTGAGACCCGCCACCCA	15222
Db	5100	GCTGGCTGGGCCAGGCTGCTTATGTAAAGGCTTCTTAAGTTTGTGTGAGACCCGCCACCCA	51599
OY	1523	CGAACCAATCAAAATAAATGACATTTCCACACCTG	1558
Db	5160	CGAACCAATCAAAATAAATGACATTTCCACACCTG	5195

RESULT 11

```

? Sequence 15682 Application US/09949016
? GENERAL INFORMATION:
? APPLICANT: VENTER, J. Craig et al
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CLO01307
? CURRENT APPLICATION NUMBER: US/09/949,016
? PRIOR FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: fastseq for Windows Version 4.0
? SEQ ID NO 15682
? LENGTH: 7198
? TYPE: DNA
? ORGANISM: Human
? S-09-949-016-15682

```


Query Match 67.1%; Score 1056; DB 7; Length 7198;
 Best Local Similarity 100.0%; Pred. No. 1.1e-22;
 Matches 1056; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 GCTATGAGAGAGTACCAATGAAAGACCTCTTACCACTGACAGCTATTGCTGCGGCC 562
 |||||||
 Db 4140 GCTATGAGAGAGTACCAATGAAAGACCTCTTACCACTGACAGCTATTGCTGCGGCC 4199

QY 563 GCGTCTGTCACCAATGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
 |||||||
 Db 4200 GCGTCTGTCACCAATGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4259

QY 623 CTGTGCTGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 682
 |||||||
 Db 4260 CTGTGCTGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 4319

QY 683 GCGCTCTGAGAGATGCGCTTCTACTCTGCGACCTCTGCGCTGCGACCGCTGCGCA 742
 |||||||
 Db 4320 GCGCTCTGAGAGATGCGCTTCTACTCTGCGACCTCTGCGCTGCGACCGCTGCGCA 4379

QY 743 CCTACCGGCGGAGCTGCGCTACCTCCCTGTAAGAGAGTGGGTTCCAGACACCTGCGCT 802
 |||||||
 Db 4380 CCTACCGGCGGAGCTGCGCTACCTCCCTGTAAGAGAGTGGGTTCCAGACACCTGCGCT 4439

QY 803 CCCCCGTTGTCGACGAGGCGCGGTAGATGACACCTGTGCGACCTGAGAGAGCCAG 862
 |||||||
 Db 4440 CCCCCGTTGTCGACGAGGCGCGGTAGATGACACCTGTGCGACCTGAGAGAGCCAG 4499

QY 863 TGCCCTCTCAGTGAAGTGTGCGGACGAGAGCTTTGTGCTAGTCTGCGACCTGCTGC 922
 |||||||
 Db 4500 TGCCCTCTCAGTGAAGTGTGCGGACGAGAGCTTTGTGCTAGTCTGCGACCTGCTGC 4559

QY 923 ACTGCGACCTGGGAGATGATGTTGTCGACCCATGCGCGCGCTGCGACGCTGCGCTCA 982
 |||||||
 Db 4560 ACTGCGACCTGGGAGATGATGTTGTCGACCCATGCGCGCGCTGCGACGCTGCGCTCA 4619

QY 983 TGCATCTGTTTACGTGCGGCGGAGTCTGTCGATGCTGCTGCGCGCTTCTCTG 1042
 |||||||
 Db 4620 TGCATCTGTTTACGTGCGGCGGAGTCTGTCGATGCTGCTGCGCGCTTCTCTG 4679

QY 1043 CCATGAGAGAGGCGAGCATATGAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTG 1102
 |||||||
 Db 4680 CCATGAGAGAGGCGAGCATATGAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTG 4739

QY 1103 TCGCCTTCGCTTGGAGCGCCAGAGATGGAAGCTGTTGTCAGAGTATGGAATTTGA 1162
 |||||||
 Db 4740 TCGCCTTCGCTTGGAGCGCCAGAGATGGAAGCTGTTGTCAGAGTATGGAATTTGA 4799

QY 1163 TGGTTAGCAGGCGCTGACAGGCGCAGTGCACCAACTACTTCTGATGCTGAGCGTT 1222
 |||||||
 Db 4800 TGGTTAGCAGGCGCTGACAGGCGCAGTGCACCAACTACTTCTGATGCTGAGCGTT 4859

QY 1223 GCGTGAAGCGCGCGCCAGCTGAGAGCCCGACAGATGCCACCGCCAGAGAGCCCTTAT 1282
 |||||||
 Db 4860 GCGTGAAGCGCGCGCCAGCTGAGAGCCCGACAGATGCCACCGCCAGAGAGCCCTTAT 4919

QY 1283 GACCCCTGGGCGCGCTGCTGCTTACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCT 1342
 |||||||
 Db 4920 GACCCCTGGGCGCGCTGCTGCTTACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCT 4979

QY 1343 GCTGACGAGCGCTGTCACAGCTCTGCTGAGGCTGAGAGAGCTCTCTGAGAGAGGTTA 1402
 |||||||
 Db 4980 GCTGACGAGCGCTGTCACAGCTCTGCTGAGGCTGAGAGAGCTCTCTGAGAGAGGTTA 5039

QY 1403 GAAAGTGAAGCTATGCTTTGGGGGAGACAGCCAGATGAAGTCTGTTGCTCAGAGACCA 1462
 |||||||
 Db 5040 GAAAGTGAAGCTATGCTTTGGGGGAGACAGCCAGATGAAGTCTGTTGCTCAGAGACCA 5099

QY 1463 GCGGCTGGGCGCGAGCTATGCTATGTAAGGCTCTCTGTTGTTGCTGAGAGAGAGAGAG 1522
 |||||||
 Db 5100 GCGGCTGGGCGCGAGCTATGCTATGTAAGGCTCTCTGTTGTTGCTGAGAGAGAGAGAG 5159

QY 1523 CGAACCAATTCCAATTAAGTACATTCACAGCTG 1558

Db 5160 CGAACCAATTCCAATTAAGTACATTCACAGCTG 5195
 |||||||

RESULT 12
 US-09-948-128-429/c
 ; Sequence 429, Application US/09948128
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig
 ; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON
 ; FILE REFERENCE: CL001294
 ; CURRENT APPLICATION NUMBER: US/09/948,128
 ; CURRENT FILING DATE: 2001-09-07
 ; NUMBER OF SEQ ID NOS: 465
 ; SEQ ID NO 429
 ; LENGTH: 5506356
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(5506356)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-948-128-429

Query Match 67.1%; Score 1056; DB 5; Length 5506356;
 Best Local Similarity 100.0%; Pred. No. 6.1e-22;
 Matches 1056; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 GCTATGAGAGAGTACCAATGAAAGACCTCTTACCACTGACAGCTATTGCTGCGGCC 562
 |||||||
 Db 4625107 GCTATGAGAGAGTACCAATGAAAGACCTCTTACCACTGACAGCTATTGCTGCGGCC 4625048

QY 563 GCGTCTGTCACCAATGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
 |||||||
 Db 4625047 GCGTCTGTCACCAATGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4624988

QY 623 CTGTGCTGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 682
 |||||||
 Db 4624987 CTGTGCTGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 4624928

QY 683 GCGCTCTGAGAGATGCGCTTCTACTCTGCGACCTCTGCGACCTGCGACCGCTGCGCA 742
 |||||||
 Db 4624927 GCGCTCTGAGAGATGCGCTTCTACTCTGCGACCTCTGCGACCTGCGACCGCTGCGCA 4624868

QY 743 CCTACCGGCGGAGCTGCGCTACCTCCCTGTAAGAGAGTGGGTTCCAGACACCTGCGCT 802
 |||||||
 Db 4624867 CCTACCGGCGGAGCTGCGCTACCTCCCTGTAAGAGAGTGGGTTCCAGACACCTGCGCT 4624808

QY 803 CCCCCGTTGTCGACGAGGCGCGGTAGATGACACCTGTGCGACCTGAGAGAGCCAG 862
 |||||||
 Db 4624807 CCCCCGTTGTCGACGAGGCGCGGTAGATGACACCTGTGCGACCTGAGAGAGCCAG 4624748

QY 863 TGCCCTCTCAGTGAAGTGTGCGGACGAGAGCTTGTGCTAGTCTGCGACCTGCTGC 922
 |||||||
 Db 4624747 TGCCCTCTCAGTGAAGTGTGCGGACGAGAGCTTGTGCTAGTCTGCGACCTGCTGC 4624688

QY 923 ACTGCGACCTGGGAGTGAAGTGTGCTGACACCTATGAGGCGCGCTGCGACCTGCGTCA 982
 |||||||
 Db 4624687 ACTGCGACCTGGGAGTGAAGTGTGCTGACACCTATGAGGCGCGCTGCGACCTGCGTCA 4624628

QY 983 TGCATCTGTTTACGTGCGGCGGAGTGTGCTGCGCAATGCTGCGCGCTTCTCTG 1042
 |||||||
 Db 4624627 TGCATCTGTTTACGTGCGGCGGAGTGTGCTGCGCAATGCTGCGCGCTTCTCTG 4624568

QY 1043 CCATGAGAGAGGCGAGCATATGAGTATGATGATGATGATGATGATGATGATGATGATGAT 1102
 |||||||
 Db 4624567 CCATGAGAGAGGCGAGCATATGAGTATGATGATGATGATGATGATGATGATGATGATGAT 4624508

QY 1103 TGCCCTTCGCTTGAAGCCAGAGATGGAAGGTTGTTGCTGAGTGAAGTGA 1162
 |||||||
 Db 4624507 TGCCCTTCGCTTGAAGCCAGAGATGGAAGGTTGTTGCTGAGTGAAGTGA 4624448

```

OY 1163 TGGTTAGCGAGCGCGTGCAGGGCCAGTGCACCCAACTACTTCTGGATGTCACGGGT 1222
DB 4624447 TGGTTAGCGAGCGCGTGCAGGGCCAGTGCACCCAACTACTTCTGGATGTCACGGGT 4624388
OY 1223 GCGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1282
DB 4624387 GCGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4624328
OY 1283 GACCCCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1342
DB 4624327 GACCCCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4624268
OY 1343 GCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1402
DB 4624267 GCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4624208
OY 1403 GAAGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1462
DB 4624207 GAAGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4624148
OY 1463 GCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1522
DB 4624147 GCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4624088
OY 1523 CGAACCAATCCAAATAAATGAATGACATTCACCGCGCTG 1558
DB 4624087 CGAACCAATCCAAATAAATGAATGACATTCACCGCGCTG 4624052

```

```

RESULT 13
US-10-144-771-2596
: Sequence 2596 Application US/10144771
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig
: TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
: FILE REFERENCE: CL001321
: CURRENT APPLICATION NUMBER: US/10/144,771
: CURRENT FILING DATE: 2002-05-15
: NUMBER OF SEQ ID NOS: 47235
: SEQ ID NO 2596
: LENGTH: 1927
: TYPE: DNA
: ORGANISM: HUMAN
: US-10-144-771-2596

```

```

Query Match 54.4% Score 856.2: DB 9: Length 1927:
Best Local Similarity 74.0%: Pred. No. 8,6e-179:
Matches 1159: Conservative 0: Mismatches 388: Indels 20: Gaps .5:

```

```

OY 1 GCGCCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
DB 365 GCGCCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 424
OY 61 AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 116
DB 425 AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 484
OY 117 GCGCGTGCAGGTTATGATCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 176
DB 485 CCGGCGACCTGGCTATGAAACAGTAGAATGCTCTCGAGGACCTGCTCCACGCGCATG 544
OY 177 GCTGCTGGTCTGCTGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 236
DB 545 AGTCTGCTGCTGCTGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 604
OY 237 TCACGTGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 296
DB 605 CCGTGTGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 664
OY 297 GCGGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 356
DB 665 GAGAACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 724

```

```

OY 357 GGTGATGCTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 416
DB 725 AGTCACTGCTCGCGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 784
OY 417 CTGGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 476
DB 785 CTGGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 844
OY 477 GCGAGCTTCTGTAACCATATGCTGCTGATGAGCGAGTGCACATGACAGCTCTCTG 536
DB 845 GCGAGCTTCTGTAACCATATGCTGCTGATGAGCGAGTGCACATGACAGCTCTCTG 904
OY 537 CAATGACAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 596
DB 905 CAATGACAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 964
OY 597 CAGCGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 656
DB 965 CAGCGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1024
OY 657 TGTGACCTGAGAGTGAAGATGATGCGCGCTGCGGCGAGATGCGCTTCACTGCGGAC 716
DB 1025 TGTGACCTGAGAGTGAAGATGATGCGCGCTGCGGCGAGATGCGCTTCACTGCGGAC 1084
OY 717 CTTCGCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 776
DB 1085 CTTCGCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1144
OY 777 AAGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 836
DB 1145 AAGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1201
OY 837 ACACCTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 896
DB 1202 ACACCTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1261
OY 897 CTTCGCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 956
DB 1262 CTTCGCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1321
OY 957 CATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1016
DB 1322 CATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1381
OY 1017 TCGCATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1076
DB 1382 TCGCATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1441
OY 1077 CCGCTGCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1136
DB 1442 CCGCTGCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1501
OY 1137 TGTGTTGAGAGTATGAGGAAATGATGCTTACGAGCGCGCGCGCGCGCGCGCGCGCG 1196
DB 1502 TGTGTTGAGAGTATGAGGAAATGATGCTTACGAGCGCGCGCGCGCGCGCGCGCGCG 1561
OY 1197 AAATGACTTGTGATGCTGACGAGTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 1256
DB 1562 AAATGACTTGTGATGCTGACGAGTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 1621
OY 1257 GATGCGCGCGCGCGAGAGCGCT---TATGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1313
DB 1622 GATGCGCGCGCGCGAGAGCGCT---TATGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1681
OY 1314 ACTTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1373
DB 1682 GAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1741
OY 1374 GTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1433
DB 1742 GTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1800
OY 1434 CCAAGATGAGTCTGGGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1484

```

Db 1801 TCTACCTTCAGTAGAGTAACATCCCGAGTAGAGCCCTGCTGGCTGGACAGATTGCATTA 1860
QY 1485 TGTAAAGCCCTTCTAGTTTCTTGAGAGACCCCGACCCGACGACCAATTCATAATAGAG 1544
Db 1861 TGAAGAACCTCCCTCCCTCTTGAAGGACCTTCCCTGGGAACCAATTCATAAAGAG 1920
QY 1545 ACATTCC 1551
Db 1921 ACTTTTC 1927

RESULT 14
US-10-338-044-2517
: Sequence 2517, Application US/10338044
: GENERAL INFORMATION:
: APPLICANT: MENDRICK, Donna
: APPLICANT: PORTER, Mark
: APPLICANT: JOHNSON, Kory
: APPLICANT: HIGGS, Brandon
: APPLICANT: CASTLE, Arthur
: APPLICANT: ELASHOFF, Michael
: TITLE OF INVENTION: Molecular Cardiototoxicology Modeling
: FILE REFERENCE: 44921-5090-01-US
: CURRENT APPLICATION NUMBER: US/10/338,044
: PRIOR APPLICATION NUMBER: US 60/303,819
: PRIOR FILING DATE: 2001-07-10
: PRIOR APPLICATION NUMBER: US 60/305,623
: PRIOR FILING DATE: 2001-07-17
: PRIOR APPLICATION NUMBER: US 60/369,351
: PRIOR FILING DATE: 2002-04-03
: PRIOR APPLICATION NUMBER: US 60/377,611
: PRIOR FILING DATE: 2002-05-06
: NUMBER OF SEQ ID NOS: 2696
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2517
: LENGTH: 2648
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. NM_133386
US-10-338-044-2517

Query Match 52.4% Score 824.8; DB 9; Length 2648;
Best Local Similarity 75.3%; Pred. No. 7.7e-172;
Matches 1084; Conservative 0; Mismatches 342; Indels 14; Gaps 4;

QY 1 GCCCCACAGCCGCGCCCTGCGAGCCGCGCTGGGACACCGATAGAGCTGAAGCAGG 60
Db 1108 GCCCCACAGCCGCGCGCTGACAGACACCGCGCTGGGACACCGATAGAGCTGACACTGG 1167
QY 61 AGCCGCCCGCAGCGGCGAGCGCCCGCC--ACAGCGCCAGGGAGCCCGCTGGCAGCGGAGCC 116
Db 1168 AGCGCGCGCTTACTCTAGAGAGCGCGCGGAGCATGCGTGGCTCTTGTCAGCGGAGCC 1227
QY 117 GCGGATCGAGGTATGATTCAGCGGGCGCGCGCGCGCGCGCTCCCGCGCGCGCGCG 176
Db 1228 CCGGAGACTGGCTATGCAACAGCAAGCTGCTCCCGGAGCTGCTCCACAGGCGCATGCGAG 1287
QY 177 CGTGTGCTGCTGCTGAACCGCGCGCGCGGCAAGGCAAGCGCTTCGAGCTTCGCGAG 236
Db 1288 AGTGTGCGCTGCTGACACCGCGCGGCTGGCAAGGCAAGCTTCGAGCTTCGAGAG 1347
QY 237 TCACTGCGACCGCTTTTGGCTGAGCTGAATCTCCTTCACGCTGCTCACTGAGCG 236
Db 1348 CCGTGTGCGCGCGCTTCTGAGAGAGGCTGAGTATCTTTAACTGATGCTCACGAGC 1407
QY 297 GCGGACCAAGCGCGCGGAGCTGCTGGGCTGAGGAGAGCTGGCGCGCTGAGCTTGT 356
Db 1408 GCGAAGCAATGCGAGGAGCTGCTGTGTGAGAGAACTGGGTACTGGATGCTTGGC 1467
QY 357 GGTATGTCTGAGAGCGGCTGATGACGAGGTGTGAAAGCGGCTCATGAGCGGCGCTGA 416

Db 1468 GGTATGTCTGAGAGCGGCTGATGACAGAGGTGTGAAATGGGCTAATGAAAGACTGA 1527
QY 417 CTGGGAGACCGCCATCCAGAGAGCCCTGTGTACCTCCAGAGGCTCTGGACAGCGCT 476
Db 1528 CTGGGAGAGTGCCTCCAGAAACCCCTATGTACCTCCCTGGAGGCTCTGGCAATCGCT 1587
QY 477 GCGAGCTTCTTGAACCATTAATGCTGTGAGAGAGGATCAGCAATGAAGCTCCGAG 536
Db 1588 GCGAGCTTCTTGAACCATTAATGCTGTGAGAGAGGATCAGCAATGAAGCTCCGAG 1647
QY 537 CAACCTGACGCTATGCTGTGCGCGCGGCTGTGTGACACCCATGAACCTGCTGTGCA 596
Db 1648 CAACCTGACGCTGTGTGTGAGCTGCGGAGCTGTGACCCATGAATGCTGTGCGCA 1707
QY 597 CAGGCTTGGGAGCTGCGCTCTTCTGTGTGCTGACAGCTGCGCTGGGAGCTTCAATGCTGA 656
Db 1708 CAGGCTTGGGAGCTGCGCTCTTCTGTGTGCTGACAGCTGCGCTGGGAGCTTCAATGCTGA 1767
QY 657 TGTGACCTGAGAGAGTGAAGTATCGGCGCTGCGGAGATGCGCTTCACTGTGGGAC 716
Db 1768 GGTGACCTTGAAGAGTGAAGTATCGGCGCTGCGGAGATGCGCTTCACTGTGGGAC 1827
QY 717 CTTCCTGCGTCTGCGAGCCCTGCGACCTACCGCGCGGAGCTGCGCTTACCTGCTAGG 776
Db 1828 CTTCCTGCGTCTGCGAGCCCTGCGACCTTACCAAGGCGCAAGCTGCGCTTACCTGCTAGG 1887
QY 777 AAGAGTGGTTCACAGACACTGCTCCCGCTTGTGTGCTCAGCAGGCGCGCTAGATGC 836
Db 1888 AAGAGTGGTTCACAGACACTGCTCCCGCTTGTGTGCTCAGCAGGCGCGCTAGATGC 1944
QY 837 ACACCTTGTGCGACCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 896
Db 1945 ATACCTGCTTCTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2004
QY 897 CTTCCTGCTGAGTCTGCGACCTGCGACCTGCGACCTGCGGAGGAGGAGGAGGAGGAGGAG 956
Db 2005 CTTCCTGCTGAGTCTGCGACCTGCGACCTGCGACCTGCGGAGGAGGAGGAGGAGGAGGAG 2064
QY 957 CAGGCGCGCTGCGACCTGCGACCTGCGACCTGCGACCTGCGACCTGCGACCTGCGACCTG 1016
Db 2065 CAGGCGCGCTGCGACCTGCGACCTGCGACCTGCGACCTGCGACCTGCGACCTGCGACCTG 2124
QY 1017 TGGCAGCTGCTGCGCGCTTCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1076
Db 2125 GCGCAGTCTGCTGCGCGCTTCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2184
QY 1077 CCGCTACTGTATATGTCGCGCGCTTCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1136
Db 2185 CCGCTACTGTATATGTCGCGCGCTTCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2244
QY 1137 TGTGTTTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1196
Db 2245 AGTGTGTTTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2304
QY 1197 AAATCTTCTGAGTGTGTCAGCGGTTGCGTGAAGCGCGCGCGCGAGCTGGAAGCGCGCGCA 1256
Db 2305 AAATCTTCTTGAAGTGTGTCAGCGGTTGCGTGAAGCGCGCGCGCGAGCTGGAAGCGCGCGCA 2364
QY 1257 GATGCGACCGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1316
Db 2365 GATGCGACCGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2424
QY 1317 TGCAGAGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1370
Db 2425 GTATGAGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2484
QY 1371 GGGGTGAGAGAGCTTCTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1430
Db 2485 GAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2543

RESULT 15

Mon May 19 10:50:38 2003

us-09-937-060a-19.rnpn

Page 14

US-09-949-016-40620

Sequence 40620, Application US/09949016

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 40620

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-40620

Job time : 4837 secs

Query Match 38.2% Score 600.6; Db 7; Length 601;
Best Local Similarity 99.8%; Pred. No. 1,4e-122;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 792 GACACCTGCTCCCGGTTGTGTCACAGGCGCGGTAGATGACACCTTGTGCACT 851
Db 1 GACACCTGCTCCCGGTTGTGTCACAGGCGCGGTAGATGACACCTTGTGCACT 60
QY 852 GGAGAGCCAGTGCCTCTCACTGACAGTGTGTCAGAGAGACTTGTGCTAGTCT 911
Db 61 GGAGAGCCAGTGCCTCTCACTGACAGTGTGTCAGAGAGACTTGTGCTAGTCT 120
QY 912 GGCACCTGCTCACTGACAGTGTGTCAGAGAGACTTGTGTCAGAGAGACTTGTG 971
Db 121 GGCACCTGCTCACTGACAGTGTGTCAGAGAGACTTGTGTCAGAGAGACTTGTG 180
QY 972 AGCTGCGTCACTGACAGTGTGTCAGAGAGACTTGTGTCAGAGAGACTTGTG 1031
Db 181 AGCTGCGTCACTGACAGTGTGTCAGAGAGACTTGTGTCAGAGAGACTTGTG 240
QY 1032 CCTCTTCCTGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1091
Db 241 CCTCTTCCTGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 1092 TGTGCGCGTGTGCGCTTCCGTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1151
Db 301 TGTGCGCGTGTGCGCTTCCGTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 1152 TGGGGAATTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1211
Db 361 TGGGGAATTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 1212 GGTTCAGGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1271
Db 421 GGTTCAGGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 1272 AGAGCCCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1331
Db 481 AGAGCCCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 1332 CCTTCCCTAG 1391
Db 541 CCTTCCCTAG 600
QY 1392 G 1392
Db 601 G 601

Search completed: May 17, 2003, 17:00:34

ORGANISM: Homo sapiens
US-09-937-886-9797

Query Match 98.5%; Score 1548.8; DB 24; Length 2494;
Best Local Similarity 99.5%; Pred. No. 2.9e-298;
Matches 1544; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1 GCCCCACAGCCGCCCTGCGACGCCGCCCTGGGAGCAGCCATTAAGAGCTGAAGGCGAG 60
DB 882 GCCCCACAGCCGCCCTGCGACGCCGCCCTGGGAGCAGCCATTAAGAGCTGAAGGCGAG 941
QY 61 AGCCGCCGCCAGCGGCGAGCG-CCCCACAGCGCGCAGGAGGCCCTGGCAGCGGCGCG 119
DB 942 AGCCGCCGCCAGCGGCGAGCGGCCCCACAGCGCGCAGGAGGCCCTGGCAGCGGCGCG 1001
QY 120 GGTGAGGTTATGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 1002 GGTGAGGTTATGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1061
QY 180 GCTGTGCTGCTGAACCCG 239
DB 1062 GCTGTGCTGCTGAACCCG 1121
QY 240 CGTGCAGCCCTTTTGGCTAGGCTGAATCTCTTCACGCTGATGCTCACTGAGCGCG 299
DB 1122 CGTGCAGCCCTTTTGGCTAGGCTGAATCTCTTCACGCTGATGCTCACTGAGCGCG 1181
QY 300 GAACCAAGCGCGGAGCTGCTGGGCTGCGAGAGCTGGGCGCGCGCGCGCGCGCGCGCG 359
DB 1182 GAACCAAGCGCGGAGCTGCTGGGCTGCGAGAGCTGGGCGCGCGCGCGCGCGCGCGCG 1241
QY 360 CATGTCTGAGAGCGGCTGATGACAGAGTGTGTAACGGGCTCATGAGAGCGGCTGACTG 419
DB 1242 CATGTCTGAGAGCGGCTGATGACAGAGTGTGTAACGGGCTCATGAGAGCGGCTGACTG 1301
QY 420 GGAAGCGCGCATCAAGAGCGCGCTGTAGCTTCCAGAGGCTCTGGCAACGCGCTGGC 479
DB 1302 GGAAGCGCGCATCAAGAGCGCGCTGTAGCTTCCAGAGGCTCTGGCAACGCGCTGGC 1361
QY 480 AGCTTCCTGAACCATATGCTGCTGATGAGAGGCTGACCAATGAACCTCTGACCAA 539
DB 1362 AGCTTCCTGAACCATATGCTGCTGATGAGAGGCTGACCAATGAACCTCTGACCAA 1421
QY 540 CTGACGCTATTGCTGTGCGCGCGCTGCTGACCCATGAACCTGCTGTCTGACAC 599
DB 1422 CTGACGCTATTGCTGTGCGCGCGCTGCTGACCCATGAACCTGCTGTCTGACAC 1481
QY 600 GCGTTGGGCGCTGCGCTTCTTCTGTGCTGACGCTGGGCTTCAATGCTGATGT 659
DB 1482 GCGTTGGGCGCTGCGCTTCTTCTGTGCTGACGCTGGGCTTCAATGCTGATGT 1541
QY 660 GGAACCTGAGAGTGAAGATATGCGGCTGCGGAGATGCGCTGCTGCTGGGACCTT 719
DB 1542 GGAACCTGAGAGTGAAGATATGCGGCTGCGGAGATGCGCTGCTGCTGGGACCTT 1601
QY 720 CTTGCTGTGCGAGCCCTGCGCACTACCGCGCGCATGCGCTTCCCTCTGAGGAG 779
DB 1602 CTTGCTGTGCGAGCCCTGCGCACTACCGCGCGCATGCGCTTCCCTCTGAGGAG 1661
QY 780 AGTGGGTTTCAAGACACTGCTTCCCGCTTGGTGTGCTGACAGGCGCGGTAATGACA 839
DB 1662 AGTGGGTTTCAAGACACTGCTTCCCGCTTGGTGTGCTGACAGGCGCGGTAATGACA 1721
QY 840 CTTTGTGCACTGGAGAGCAGTGCCTTCACTGAGAGTGTGCTGCGGAGGACTT 899
DB 1722 CTTTGTGCACTGGAGAGCAGTGCCTTCACTGAGAGTGTGCTGCGGAGGACTT 1781
QY 900 TGTGCTAGTCTGGCACTGCTGCACTGCACTGCGAGTGAATGTTTCTGCAACCAT 959
DB 1782 TGTGCTAGTCTGGCACTGCTGCACTGCACTGCGAGTGAATGTTTCTGCAACCAT 1841
QY 960 GGGCGGCTGTGAGCTGCTGATGATCTGTCTAGCTGCGGCGGAGTGTCTGCTGC 1019
DB 1019 GGGCGGCTGTGAGCTGCTGATGATCTGTCTAGCTGCGGCGGAGTGTCTGCTGC 1019

DB 1842 GGGCGGCTGTGAGCTGCTGATGATCTGTCTAGCTGCGGCGGAGTGTCTGCTGC 1901
QY 1020 CATGCTGCTGGGCTCTTCCGCGCATGGAAGGAGGAGGATTAAGATGAATGCC 1079
DB 1902 CATGCTGCTGGGCTCTTCCGCGCATGGAAGGAGGAGGATTAAGATGAATGCC 1961
QY 1080 CTACTTGTATATGTCGCCGCTGTGCGCTTCCGCTTGGAGCCCAAGATGGAAGTGT 1139
DB 1962 CTACTTGTATATGTCGCCGCTGTGCGCTTCCGCTTGGAGCCCAAGATGGAAGTGT 2021
QY 1140 GTTTCAGTGGAGGGAATGATGTTAGCGAGCGCTGACGAGGCGAGGTGACCCAAA 1199
DB 2022 GTTTCAGTGGAGGGAATGATGTTAGCGAGCGCTGACGAGGCGAGGTGACCCAAA 2081
QY 1200 CTACTTGTATGTTGAGGCTTCCGCTGGAAGCCCGCCCAAGCTGGAAGCCCAAGAT 1259
DB 2082 CTACTTGTATGTTGAGGCTTCCGCTGGAAGCCCGCCCAAGCTGGAAGCCCAAGAT 2141
QY 1260 GCCACGCCAGAAAGGCGCTTATGACCCCTGGGCGCGCTGTGCTTACTTGC 1319
DB 2142 GCCACGCCAGAAAGGCGCTTATGACCCCTGGGCGCGCTGTGCTTACTTGC 2201
QY 1320 AGGACCTTCTCTCTTCCCTGAGGCTGAGGCGCTGTCACAGCTCTGTGGGCTGAG 1379
DB 2202 AGGACCTTCTCTCTTCCCTGAGGCTGAGGCGCTGTCACAGCTCTGTGGGCTGAG 2261
QY 1380 GAGACTCTCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAG 1439
DB 2262 GAGACTCTCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAG 2321
QY 1440 TGAAGTCTGCTGAGAGGCGCAGCTGCTGGGCGCGCTGCTGCTTATGAGGCTTCTAG 1499
DB 2322 TGAAGTCTGCTGAGAGGCGCAGCTGCTGGGCGCGCTGCTGCTTATGAGGCTTCTAG 2381
QY 1500 TTTGTTCTGAGAGCCCGCCAGCCGAGAACCAATCAATTAAGTGAATGCCAGCTGA 1559
DB 2382 TTTGTTCTGAGAGCCCGCCAGCCGAGAACCAATCAATTAAGTGAATGCCAGCTGA 2441
QY 1560 AAAAAAAAAA 1571
DB 2442 CACCAACATGAA 2453

RESULT 8
US-09-649-163-9696
; Sequence 9696, Application US/09649163
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Galvin, Katherine A.
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kingsbury, Gillian A.
; APPLICANT: Welch, Nadine S.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Williamson, Mark
; APPLICANT: Richardson, Jennifer
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Fraser, Christopher C.
; APPLICANT: Villalva, Jean-Luc M.G.
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: Siles-Santillana, Immaculada
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; APPLICANT: Busfield, Samantha J.
; APPLICANT: Deeds, James
; APPLICANT: Lee, John
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1164-001
; CURRENT APPLICATION NUMBER: US/09/649,163
; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/150,608
; PRIOR FILING DATE: 1999-08-25


```

RESULT 5
PCT-US02-24623-1
; Sequence 1, Application PC/TUS0224623
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SPKs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-104
; CURRENT APPLICATION NUMBER: PCT/US02/24623
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/310,362
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/357,501
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-24623-1

```

```

Query Match          98.6%; Score 1550.8; DB 1; Length 1783;
Best Local Similarity 99.5%; Pred. No. 1.1e-298;
Matches 1566; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```

```

QY 1 GCGCCACAGCGGCGCTGGAGCGCCGCTGGGACAGCAGCAGATAGAGAGCTGAAGCAG 60
DB 207 GCGCCACAGCGGCGCTGGAGCGCCGCTGGGACAGCAGCAGATAGAGAGCTGAAGCAG 266
QY 61 AGCGCGCGCGCGAGCGGAGCGG-CCCGACAGCGCGGAGCGGAGCGGCGGAGCGCGG 119
DB 267 AGCGCGCGCGCGAGCGGAGCGGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGCGG 326
QY 120 GCGCGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 179
DB 327 GCGCGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 386
QY 180 GCGGAGCTGCTGAACCGCGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCT 239
DB 387 GCGGAGCTGCTGAACCGCGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCT 446
QY 240 CGTGACGCGCGCTTTGGCTGAGGCTGAATCTCTTCACGCTGATGCTCACTGAGCGG 299
DB 447 CGTGACGCGCGCTTTGGCTGAGGCTGAATCTCTTCACGCTGATGCTCACTGAGCGG 506
QY 300 GAACCGAGCGGCGGAGCGGAGCGGCTGGGCTGAGAGAGCTGGGCGCTGGAGCGCT 359
DB 507 GAACCGAGCGGCGGAGCGGAGCGGCTGGGCTGAGAGAGCTGGGCGCTGGAGCGCT 566
QY 360 CATGCTGAGAGAGGCGCTGATGACAGAGGTGAGAGAGGCTGATGAGAGGCGCTGACT 419
DB 567 CATGCTGAGAGAGGCGCTGATGACAGAGGTGAGAGAGGCTGATGAGAGGCGCTGACT 626
QY 420 GAGAGCGCGCATCAAGAGCGCGCTGTGTAGCGCTCCAGCAGGAGCTGTGGCAAGCGG 479
DB 627 GAGAGCGCGCATCAAGAGCGCGCTGTGTAGCGCTCCAGCAGGAGCTGTGGCAAGCGG 686
QY 480 AGCTTCCTTGAACCATTAATGCTGGCTATGAGAGGTGACCAATGAAGAGCTCTGACCA 539
DB 687 AGCTTCCTTGAACCATTAATGCTGGCTATGAGAGGTGACCAATGAAGAGCTCTGACCA 746
QY 540 CTGACAGCTATGCTGCTGCGCGCGGCTGCTGATGACCATGACCTGCTGCTGACAG 599
DB 747 CTGACAGCTATGCTGCTGCGCGCGGCTGCTGATGACCATGACCTGCTGCTGACAG 806
QY 600 GCGTTCGGGCGCTGCGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
DB 807 GCGTTCGGGCGCTGCGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
QY 660 GAGACCTAGAGAGTGAAGATGAGGCTGAGGAGAGTGAAGAGTGAAGAGTGAAGAGT 719
DB 867 GAGACCTAGAGAGTGAAGATGAGGCTGAGGAGAGTGAAGAGTGAAGAGTGAAGAGT 926

```

```

QY 720 CCGCGCTGAGAGAGCGCGCTGAGAGCTACCGCGGCGGAGTGGCTACCTGCTAGAGAG 779
DB 927 CCGCGCTGAGAGAGCGCGCTGAGAGCTACCGCGGCGGAGTGGCTACCTGCTAGAGAG 986
QY 780 AGTGGCTTCCAGAGAGCTGCGCTCCCGCTTGTGTGCTGAGAGAGGCGCGGTAGATG 839
DB 987 AGTGGCTTCCAGAGAGCTGCGCTCCCGCTTGTGTGCTGAGAGAGGCGCGGTAGATG 1046
QY 840 CCTGTGACAGTGGAGAGCGAGTGGCGCTGACAGAGAGTGGCGGAGAGAGCT 899
DB 1047 CCTGTGACAGTGGAGAGCGAGTGGCGCTGACAGAGAGTGGCGGAGAGAGCT 1106
QY 900 TGTGTAGTCTGAGAGAGCTGAGAGAGTGGCGCTGAGAGAGTGGCGGAGAGAGCT 959
DB 1107 TGTGTAGTCTGAGAGAGCTGAGAGAGTGGCGCTGAGAGAGTGGCGGAGAGAGCT 1166
QY 960 GGGCGCGCTGAGAGAGCTGAGAGAGTGGCGCTGAGAGAGTGGCGGAGAGAGCT 1019
DB 1167 GGGCGCGCTGAGAGAGCTGAGAGAGTGGCGCTGAGAGAGTGGCGGAGAGAGCT 1226
QY 1020 CATGCTGCGCGCGCTGCTGCTGCGCATGAGAGAGGCGAGCATATGAGATGAATGCC 1079
DB 1227 CATGCTGCGCGCGCTGCTGCTGCGCATGAGAGAGGCGAGCATATGAGATGAATGCC 1286
QY 1080 CTACTGTATATATGCTGCGCGCTGCTGCGCATGAGAGAGGCGAGATGAGAGAGT 1139
DB 1287 CTACTGTATATATGCTGCGCGCTGCTGCGCATGAGAGAGGCGAGATGAGAGAGT 1346
QY 1140 GTTGTGAGTATGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 1199
DB 1347 GTTGTGAGTATGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 1406
QY 1200 CTACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
DB 1407 CTACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1466
QY 1260 GCGACCGCGAGAGAGAGCTTATGACCGCTGAGCGCGGCTGCTTATGATGATGAT 1319
DB 1467 GCGACCGCGAGAGAGAGCTTATGACCGCTGAGCGCGGCTGCTTATGATGATGAT 1526
QY 1320 AGGAGCGCTTCTCTCTTCCCTAGGAGCTGAGAGGCTGCTGAGAGGCTGCTGAG 1379
DB 1527 AGGAGCGCTTCTCTCTTCCCTAGGAGCTGAGAGGCTGCTGAGAGGCTGCTGAG 1586
QY 1380 GAGAGCTCTGAGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1439
DB 1587 GAGAGCTCTGAGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1646
QY 1440 TGAAGTCTGAGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1499
DB 1647 TGAAGTCTGAGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1706
QY 1500 TTTGTGCTGAGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1559
DB 1707 TTTGTGCTGAGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1766
QY 1560 AAAAAAAAAAAAAA 1573
DB 1767 AAAAAAAAAAAAAA 1780

```

```

RESULT 6
US-09-617-081-2136
; Sequence 2136, Application US/09617081
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE REFERENCE: 1600.1146-001
; CURRENT APPLICATION NUMBER: US/09/617,081
; PRIOR APPLICATION NUMBER: 60/143,929

```


OY 1500 TTTTGTCTGAGACCCCGACCCGACCAATTCGAAATTAAGTGCATTCGCCGCTGA 1559
 |||||||
 Db 2382 TTTTGTCTGAGACCCCGACCCGACCAATTCGAAATTAAGTGCATTCGCCGCTGA 2441
 OY 1560 AAAAAAAAAAAAAA 1573
 | |||||||
 Db 2442 ACCAAAAAAAAAAAAA 2455

RESULT 4
 US-60-278-16055
 : Sequence 16055, Application US/60278258
 : GENERAL INFORMATION:
 : APPLICANT: Morris, MacDonald
 : APPLICANT: Lal, Preeti
 : APPLICANT: Diep, Binh
 : TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
 : TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
 : FILE REFERENCE: GX-0010-1 P
 : CURRENT APPLICATION NUMBER: US/60/278,258
 : NUMBER OF SEQ ID NOS: 2001-03-23
 : SOFTWARE: PERL Program
 : SEQ ID NO 16055
 : LENGTH: 1991
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc.feature
 : OTHER INFORMATION: Inocyte ID No: 221818.6
 US-60-278-258-16055

Query Match 99.0% Score 1557; DB 71; Length 1991;
 Best Local Similarity 99.9% Pred. No. 6.5e-300;
 Matches 1568; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 GCCCCACAGCCGCCCTCTGAGAGCCCGCTGGGAGACCGATAGAGAGCTGAAGGAGG 60
 Db 423 GCCCCACAGCCGCCCTCTGAGAGCCCGCTGGGAGACCGATAGAGAGCTGAAGGAGG 482
 OY 61 AGCCGCGCGGACGAGGAGAGG--CCCCACAGGCGCAGGAGCCCTGGCAGGAGGAGCGG 119
 |||||||
 Db 483 AGCCGCGCGGACGAGGAGAGG--CCCCACAGGCGCAGGAGCCCTGGCAGGAGGAGCGG 542
 OY 120 GGTGAGAGTTATGATCCAGGCGGCGGCGCCCGAGGCGTGTCCCGCGCCCTGCGCGGT 179
 |||||||
 Db 543 GGTGAGAGTTATGATCCAGGCGGCGGCGCCCGAGGCGTGTCCCGCGCCCTGCGCGGT 602
 OY 180 GGTGAGAGTTATGATCCAGGCGGCGGCGCCCGAGGCGTGTCCCGCGCCCTGCGCGGT 239
 |||||||
 Db 603 GGTGAGAGTTATGATCCAGGCGGCGGCGCCCGAGGCGTGTCCCGCGCCCTGCGCGGT 662
 OY 240 GGTGAGAGTTATGATCCAGGCGGCGGCGCCCGAGGCGTGTCCCGCGCCCTGCGCGGT 299
 |||||||
 Db 663 GGTGAGAGTTATGATCCAGGCGGCGGCGCCCGAGGCGTGTCCCGCGCCCTGCGCGGT 722
 OY 300 GAACACAGGCGGAGCTGTGCGGTGAGAGAGCTGTGCGCGCCCTGCGCGGTCTGTGCT 359
 |||||||
 Db 723 GAACACAGGCGGAGCTGTGCGGTGAGAGAGCTGTGCGCGCCCTGCGCGGTCTGTGCT 782
 OY 360 CATGTCTGAGAGCGGAGCTGTGAGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTGAG 419
 |||||||
 Db 783 CATGTCTGAGAGCGGAGCTGTGAGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTGAG 842
 OY 420 GAGAGCGGCGCATTCAGAGAGCCCTGTGTAGCTTCCAGAGAGCTGTGAGAGAGCTGTG 479
 |||||||
 Db 843 GAGAGCGGCGCATTCAGAGAGCCCTGTGTAGCTTCCAGAGAGCTGTGAGAGAGCTGTG 902
 OY 480 AGCTCTCTTGAACCATTTATGCGGGCTATGAGCAGAGCTACCATTAAGAGAGCTGTGAG 539
 |||||||
 Db 903 AGCTCTCTTGAACCATTTATGCGGGCTATGAGCAGAGCTACCATTAAGAGAGCTGTGAG 962

OY 540 CTGACAGCTATTGCTGTGCGGCGGCGTGTGATCCCATGAAACGCTGTCTGACAC 599
 |||||||
 Db 963 CTGACAGCTATTGCTGTGCGGCGGCGTGTGATCCCATGAAACGCTGTCTGACAC 1022
 OY 600 GCGTGGGGGCTGGCGCTCTTCTGTGCTACCGCTGGGCTTCAATGTGATCT 659
 |||||||
 Db 1023 GCGTGGGGGCTGGCGCTCTTCTGTGCTACCGCTGGGCTTCAATGTGATCT 1082
 OY 660 GGAAGCTAGAGAGTGAAGATATGCGGCTGTGGGAGATGCGCTTCACTGTGGCACC 719
 |||||||
 Db 1083 GGAAGCTAGAGAGTGAAGATATGCGGCTGTGGGAGATGCGCTTCACTGTGGCACC 1142
 OY 720 CCGTGGCTGGAGCCCTGCGCACCTACCGCGCCGACGTGAGCTTCACTGTGAGAG 779
 |||||||
 Db 1143 CCGTGGCTGGAGCCCTGCGCACCTACCGCGCCGACGTGAGCTTCACTGTGAGAG 1202
 OY 780 AGTGGTTTCAAGACACCTGCTCCCGCTGTGTGTCCAGCAGGCGCCGATAGATGACA 839
 |||||||
 Db 1203 AGTGGTTTCAAGACACCTGCTCCCGCTGTGTGTCCAGCAGGCGCCGATAGATGACA 1262
 OY 840 CCTTGTGACACAGAGAGAGCCAGTGCCTCTCACTGAGACAGTGGTCCGACGAGACT 899
 |||||||
 Db 1263 CCTTGTGACACAGAGAGAGCCAGTGCCTCTCACTGAGACAGTGGTCCGACGAGACT 1322
 OY 900 TGTGCTAGTCTGGGACGCTGGACCTGGACCTGGGAGTGAATGTTGCTGACCCAT 959
 |||||||
 Db 1323 TGTGCTAGTCTGGGACGCTGGACCTGGACCTGGGAGTGAATGTTGCTGACCCAT 1382
 OY 960 GGGCGGCTGTGAGCTGGGCTTATGATCTGTCTACGTCGGCGGGAGAGTGTCTGTC 1019
 |||||||
 Db 1383 GGGCGGCTGTGAGCTGGGCTTATGATCTGTCTACGTCGGCGGGAGAGTGTCTGTC 1442
 OY 1020 CATGCTGTGCGGCTCTTCTGCGCATGAGAAAGGCGAGCATATGAGATGATGCC 1079
 |||||||
 Db 1443 CATGCTGTGCGGCTCTTCTGCGCATGAGAAAGGCGAGCATATGAGATGATGCC 1502
 OY 1080 CTACTTGTATATGTGCGCCGCTGTGCGCTTCCGCTTGGAGCCCAAGATGGGAAAGGT 1139
 |||||||
 Db 1503 CTACTTGTATATGTGCGCCGCTGTGCGCTTCCGCTTGGAGCCCAAGATGGGAAAGGT 1562
 OY 1140 GTTGTGAGAGTATGAGGAAATGATGATGAGAGGCGTGAAGGCGAGTGCACCCAAA 1199
 |||||||
 Db 1563 GTTGTGAGAGTATGAGGAAATGATGATGAGAGGCGTGAAGGCGAGTGCACCCAAA 1622
 OY 1200 CTACTTGTGATGTGACAGGCTGTGCGAGAGCCCGGACCTGTGAAGCCCGACGAGAT 1259
 |||||||
 Db 1623 CTACTTGTGATGTGACAGGCTGTGCGAGAGCCCGGACCTGTGAAGCCCGACGAGAT 1682
 OY 1260 GCCACCGCCAGAGAGCCCTTATGACCCCTGGGCGCGGCTGTGAGCTTACTACTGCG 1319
 |||||||
 Db 1683 GCCACCGCCAGAGAGCCCTTATGACCCCTGGGCGCGGCTGTGAGCTTACTACTGCG 1742
 OY 1320 AGGACCTCTTCTCTTCTGAGAGCTGTGAGAGGCTGTGACACAGCTTCTGTGGGGTGA 1379
 |||||||
 Db 1743 AGGACCTCTTCTCTTCTGAGAGCTGTGAGAGGCTGTGACACAGCTTCTGTGGGGTGA 1802
 OY 1380 GAGACTCTCTGAGAGAGGCTGTGAGAGGCTGTGAGAGCTTCTTGTGGGGGACAGCCAGA 1439
 |||||||
 Db 1803 GAGACTCTCTGAGAGAGGCTGTGAGAGGCTGTGAGAGCTTCTTGTGGGGGACAGCCAGA 1499
 OY 1440 TGAAGTCTGTGGGTGAGAGCCAGCTGTGGGCGCAGCTGTGATGAGAGCCCTCTAG 1499
 |||||||
 Db 1863 TGAAGTCTGTGGGTGAGAGCCAGCTGTGGGCGCAGCTGTGATGAGAGCCCTCTAG 1922
 OY 1500 TTTGTCTGAGAGCCCGACCCGACGAAACCAATTCGAAATTAAGTGAATTCGAGCTGA 1559
 |||||||
 Db 1923 TTTGTCTGAGAGCCCGACCCGACGAAACCAATTCGAAATTAAGTGAATTCGAGCTGA 1982
 OY 1560 AAAAAAAAAA 1568
 | |||||||
 Db 1983 AAAAAAAAAA 1991

```

Db      1381  AGACTCCTCTGGAGAAAGGCTGANAAGCTGGAGGCTATGCTTTGGGGGGGACAGCCAGAT 1440
              |||
QY      1441  GAACTCCTGGGCTAGAGAGCCAGCTGGCTGGGSCCCAGCTGCTATGTAAGGCTTTAGT 1500
              |||
Db      1441  GAACTCCTGGGCTAGAGAGCCAGCTGGCTGGGSCCCAGCTGCTATGTAAGGCTTTAGT 1500
              |||
QY      1501  TTGTCTTGAANCCCCCACCAGCCAGAACCAATTCAAATTAAGTGACATTTCCAGCTGAA 1560
              |||
Db      1501  TTGTCTTGAANCCCCCACCAGCCAGAACCAATTCAAATTAAGTGACATTTCCAGCTGAA 1560
              |||
QY      1561  AAAAAAAAAAAAAA 1573
              |||
Db      1561  AAAAAAAAAAAAAA 1573
              |||

RESULT 3
US-09-526-996-1216
; Sequence 1216, Application US/09526996
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM
; TITLE OF INVENTION: HUMAN PROSTATE AND BREAST LIBRARIES
; PRT: REFERENCE, 1600, 1098, -001

```

```

CURRENT FILING DATE: 2000-03-15
EARLIER APPLICATION NUMBER: 60/125,469
EARLIER FILING DATE: 1999-03-22
EARLIER APPLICATION NUMBER: 60/125,560
EARLIER FILING DATE: 1999-03-22
NUMBER OF SEQ ID NOS: 1296
SOFTWARE: FastsEQ for Windows Version 3.0
SEQ ID NO 1216
LENGTH: 2494
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(2494)
OTHER INFORMATION: n - A,T,C or G
US-09-526-926-1216

```

Query Match	99.0%;	Score 1557.2;	DB 19;	Length 2494;
Best Local Similarity	99.7%;	Pred. No. 6.1e-300;		
Matches 1570;	Conservative	0;	Mismatches 3;	Indels 1;
			Gaps	1

Oy	1	GCCTCAGACGGGGCCCTGGAGAGCCCGCCCTGGGCACACCGATTAAGAGAGCTGAAGCAGG	60
Dd	882	GCCCCACAGCCGGGCCCTGGCAGCCCGCCCTGGGCACACCGATTAAGAGAGCTGAAGCAGG	941
Oy	61	AGCCCGCGCCACAGGGCAGCG -CCCCACAGCGCCAGGAGACCCCTGGCAGCGGAGCCAGG	119
Dd	942	AGCCCGCGCCACAGGGCAGCGCCCCACAGCGCCAGGGAGACCCCTGGCAGCGGAGCCAGG	1001
Oy	120	GCTCGAGGTTATGGATCCAGCGGGGCGGCCCCCGGGCGTGCTCCCGAGCCCTGCAGCGT	179
Dd	1002	GCTCGAGGTTATGGATCCAGCGGGGCGGCCCCCGGGCGTGCTCCCGAGCCCTGCAGCGT	1061
Oy	180	GCTGCTCTCTGTGAACCCCGCGCGGGCGGCAAGGGCAAGGCGCTTGCAGCTTTCGGAGTCA	239
Dd	1062	GCTGCTCTCTGTGAACCCCGCGCGGGCGGCAAGGGCAAGGCGCTTGCAGCTTTCGGAGTCA	1121
Oy	240	CGTCAGACCCCTTTTGGCTCAGGCTGAAATCTCTCTCACGCTGATGCTACTGAGCGGCG	299
Dd	1122	CGTCAGACCCCTTTTGGCTCAGGCTGAAATCTCTCTCACGCTGATGCTACTGAGCGGCG	1181
Oy	300	GAACCCAGCGCGGAGCGTGGTGGCTGGAGAAGCTGGGCCCGCTGGAGCGCTTGGTGT	359
Dd	1182	GAACCCAGCGCGGAGCGTGGTGGCTGGAGAAGCTGGGCCCGCTGGAGCGCTTGGTGT	1241
Oy	360	CATGCTTCGAGACGGGCTGATGCACAGAGTGGTGAACCGGGCTCATGAGCGGCTGACTG	419
Dd	1242	CATGCTTCGAGACGGGCTGATGCACAGAGTGGTGAACCGGGCTCATGAGCGGCTGACTG	1301

Oy	420	GGAGACGGCCATCCAGAAAGCCCTGTGTAGCTCCAGAGGCTGTGGCAACGGCGTGGC	479
Db	1302	GGAGACGGCCATCCAGAAAGCCCTGTGTAGCCCTCCAGAGGCTGTGGCAACGGCGTGGC	1361
Oy	480	AGCTTCCTTGAAACCATTAATCTGGCTATAGAGAGGTCACCATGTGAAGACCTTCGACCA	539
Db	1382	AGCTTCCTTGAAACCATTAATCTGGCTATAGAGAGGTCACCAATGTGAAGACCTTCGACCA	1421
Oy	540	CTGCACGCTATTGCTGTGCCGCGGCTGTGCACCATGAACCTGTGTCTTCGACAC	599
Db	1422	CTGCACGCTATTGCTGTGCCGCGGCTGTGTGCACCATGAACCTGTGTCTTCGACAC	1481
Oy	600	GGCTTCGGGAGCTGGCGCTCTTCTGTGTGCTCAGCTGTGGCTTGGGCTTCATTGCTGATGT	659
Db	1482	GGCTTCGGGAGCTGGCGCTCTTCTGTGTGCTCAGCTGTGGCTTGCATTGCTGATGT	1541
Oy	660	GGACCTTAGAGTAGAAGATATCGGGGTCTGGGGAGATAGGGCTTCACCTGTGGCACCTT	719
Db	1542	GGACCTTAGAGTAGAAGATATCGGGGTCTGGGGAGATAGGGCTTCACCTGTGGCACCTT	1601
Oy	720	CCCTGCTGTGACAGCCCTGTGGCACCTACCGCGGCGCACCTAGGCTCACTCCCTGTAGAG	779
Db	1602	CCTGCTGTGACAGCCCTGTGGCACCTACCGCGGCGCACCTAGGCTCACTCCCTGTAGAG	1661
Oy	780	AGTGGGTTCCAAAGACACCTGCTCCCGGTTGTGTCCAGAGGGCCGGTAGATGCACA	839
Db	1662	AGTGGGTTCCAAAGACACCTGCTCCCGGTTGTGTCCAGAGGGCCGGTAGATGCACA	1721
Oy	840	CGTTGTGCACTGGAGAGGACATGGCCCTCTCACTGGAAAGTGTGTCCGAGAGAGACTT	899
Db	1722	CGTTGTGCACTGGAGAGGACATGGCCCTCTCACTGGAAAGTGTGTCCGAGAGAGACTT	1781
Oy	900	TGTGCTAGTCCCTGTGCACCTGTGTGCACCTGTGCACCTGTGAATGTTTGTGTGCACCAT	959
Db	1782	TGTGCTAGTCCCTGTGCACCTGTGTGCACCTGTGCACCTGTGAATGTTTGTGTGCACCAT	1841
Oy	960	GGGCGCGTGTGACAGCTGGCGCTCATGTGTTCTACGTGGCGGGAGAGTGTCTGTGC	1019
Db	1842	GGGCGCGTGTGACAGCTGGCGCTCATGTGTTCTACGTGGCGGGAGAGTGTCTGTGC	1901
Oy	1020	CATGCTGTGTGGCTCTTCTGTGGCATGGAAGGGCAGGCTATGGAATGATGCTGCC	1079
Db	1902	CATGCTGTGTGGCTCTTCTGTGGCATGGAAGGGCAGGCTATGGAATGATGCTGCC	1961
Oy	1080	CTACTTGGTATATGTGCCCGGCGGCTTCGCTTGAAGGCCCAAGATGGGAAGGTGT	1139
Db	1962	CTACTTGGTATATGTGCCCGGCGGCTTCGCTTGAAGGCCCAAGATGGGAAGGTGT	2021
Oy	1140	GTTTGTGCAATGATGGGAATTTGATGTGTTAGCGAGGCCCTGCAAGGGCCAGGTGCACCCAA	1199
Db	2022	GTTTGTGCAATGATGGGAATTTGATGTGTTAGCGAGGCCCTGCAAGGGCCAGGTGCACCCAA	2081
Oy	1200	CTACTTCTGGATGCTCAGCGGTTGCGTGGAGCCCGCCAGCTGGAAAGCCCCAGCAGAT	1259
Db	2082	CTACTTCTGGATGCTCAGCGGTTGCGTGGAGCCCGCCAGCTGGAAAGCCCCAGCAGAT	2141
Oy	1280	GCCACCGCCAAAGAGCCCTTATGACCCCTGGGCGCGCGTGTGCCCTTATGTCTACTTGC	1319
Db	2142	GCCACCGCCAAAGAGCCCTTATGACCCCTGGGCGCGCGTGTGCCCTTATGTCTACTTGC	2201
Oy	1320	AGGACCCCTTCTCTTCCCTTAGGGCTGACAGGCTCTTCCACAGCTCTGTGGGGTGTAG	1379
Db	2202	AGGACCCCTTCTCTTCCCTTAGGGCTGACAGGCTCTTCCACAGCTCTGTGGGGTGTAG	2261
Oy	1380	GAGACTCTCTGGAAGAGGTTGAAGAGTGAAGCTATCTCTTTGGGGGACAGGCCAGAA	1439
Db	2262	GAGACTCTCTGGAAGAGGTTGAAGAGTGAAGCTATCTTTGGGGGACAGGCCAGAA	2321
Oy	1440	TGAAGTCTGTGGCTAGAGAGCCCAAGCTGTGGGCGCCAGCTGTGCTATGTAAAGCCCTTAC	1499
Db	2322	TGAAGTCTGTGGCTAGAGAGCCCAAGCTGTGGGCGCCAGCTGTGCTATGTAAAGCCCTTAC	2381

Db	1962	CTACTGGTATATGTGCCCGGTGGTCCTTCCTTGAGGCCAAGGATGGGAAGGTGT	2021
Qy	1140	GTTTTCAGTGGATGGGGAATTTGATGGTTAGCCAGAGCCGTCGAGGGCCAGGTACACCCAAA	1199
Db	2022	CTTTTGCAGTGGATGGGGAATTTGATGGTTAGCCAGAGCCGTCGAGGGCCAGGTACACCCAAA	2081
Qy	1200	CTACTTCGATGTGTACAGCGGTTGGCGTGGAGGCCCGCCAGCTGGAAAGCCCCAGCAGAT	1259
Db	2082	CTACTTCGATGGTTCAGCGGTTGGCGTGGAGGCCCGCCAGCTGGAAAGCCCCAGCAGAT	2141
Qy	1260	GCCACCGGCAGAGAGCCCTTATGACCCCTGGGCGCGGCTGTGCTTACTGTCTACCTTC	1319
Db	2142	GCCACCGGCAGAGAGCCCTTATGACCCCTGGGCGCGGCTGTGCTTACTGTCTACCTTC	2201
Qy	1320	AGGACCTTCCTCCTCCTTCCCTAAGGCGTCGAGGCGCTCTCCACAGCTCTGTGGGGGTGGAG	1379
Db	2202	AGGACCTTCCTCCTCCTTCCCTAAGGCGTCGAGGCGCTCTCCACAGCTCTGTGGGGGTGGAG	2261
Qy	1380	GAGACTCCTCTGGGGAAGGGTGAAGAGTGAAGGTATGCTTTGGGGGGAGACGGCCAAA	1439
Db	2262	GAGACTCCTCTGGGGAAGGGTGAAGAGTGAAGGTATGCTTTGGGGGGAGACGGCCAAA	2321
Qy	1440	TGAAGTCCCTGGGTGAGAGAGCCAGACGCTGGGCGCCAGCGCTATGTAAGGCCCTTTAG	1499
Db	2322	TGAAGTCCCTGGGTGAGAGAGCCAGACGCTGGGCGCCAGCGCTATGTAAGGCCCTTTAG	2381

QY	1261	CGACGCGGAAGACCCCTTATGACCCCTGCGCGCGCTGAGCTTATGCTGACTACTGCA	1320
Db	1261	CGACGCCGGAAGACCCCTTATGACCCCTGCGCGCGCTGAGCTTATGCTGACTACTGCA	1320
QY	1321	GGACCCCTCCCTCCCTTCCCTTAAAGCGTCGACGGCCCTGTCCACAGCTCCCTGTGTGGGGTGGAGG	1380
Db	1321	GGACCCCTCCCTCCCTTCCCTTAAAGGGTGTGAGGGCCCTGTCCACAGCTCCCTGTGTGGGGTGGAGG	1380
QY	1381	AGCACTCCCTCGGAGAGAGGTGTGACAGAGTGGAGGCTATGCTTTTGGGGGAGCAGGCCAAGT	1440
Db	1381	AGCACTCCCTCGGAGAGAGGTGTGACAGAGTGTAGAGCTATGCTTTTGGGGGAGCAGGCCAAGT	1440
QY	1441	GAAGTCCCTGGGTGACGAGCCCAAGCTGCTGTGGGCCCAAGCTGCTTATGTAAAGCCCTTCTGTAGT	1500
Db	1441	GAAGTCCCTGGGTGACGAGCCCAAGCTGCTGTGGGCCCAAGCTGCTTATGTAAAGCCCTTCTGTAGT	1500
QY	1501	TTGTTCCTGAGACCCCAACCCCAAGCAACAATTCAAATTAAGGTGACATTTCCAGCCGTGAA	1560
Db	1501	TTGTTCCTGAGACCCCAACCCCAAGCAACAATTCAAATTAAGGTGACATTTCCAGCCGTGAA	1560
QY	1561	AAAAAAAAAAAAAAAA 1573	
Db	1561	AAAAAAAAAAAAAAAA 1573	

QY	301	ATGACGGCGGGAGACTGTGTCGGGTGCGAGAGAGACTGGGCGGTGGAGAGCGCTGTGATGTC	360
Dp	301	AACCAAGCGCGGAGACTGTGTCGGGTCCGAGAGAGACTGGGCGGTGGAGAGCGCTGTGATGTC	360
QY	361	ATGTCTGAGAGACGGGCTGATGCACAGAGCTGTGAACGGGCTCATGAGACGGGCTGAACTGG	420
Dp	361	ATGTCTGAGAGACGGGCTGATGCACAGAGCTGTGAACGGGCTCATGAGACGGGCTGAACTGG	420
QY	421	GAAACCGGCATCCAAACCCCTGTGATGCTCCGACGAGGCTGAGCAAGGCGTGGCA	480
Dp	421	GAAACCGGCATCCAAACCCCTGTGATGCTCCGACGAGGCTGAGCAAGGCGTGGCA	480
QY	481	GCTTCTTCAACCATTTATGTCGGCTATGAGACAGTCAACCAATGAAGAGACTCTGACCAAC	540
Dp	481	GCTTCTTCAACCATTTATGTCGGCTATGAGACAGTCAACCAATGAAGAGACTCTGACCAAC	540
QY	541	TGACACCTATTCTGTGCGCGCGGCTCTTCAACCATTAACCTGCTGTCGTCACAG	600
Dp	541	TGACACCTATTCTGTGCGCGCGGCTCTTCAACCATTAACCTGCTGTCGTCACAG	600
QY	601	GCTTTCGGGCTCGGCTCTTCTGTGTCGTCAGGCTGGCTGGGCTTCAATTCGATGTG	660
Dp	601	GCTTTCGGGCTCGGCTCTTCTGTGTCGTCAGGCTGGCTGGGCTTCAATTCGATGTG	660

```

RESULT 2
US-60-125-593-12
: Sequence 12, Application US/60125593
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Tang, Y. Tom
: APPLICANT: Yue, Henry
: APPLICANT: Baughn, Mariah R.
: APPLICANT: Azimzal, Yalala
: TITLE OF INVENTION: REGULATORS OF INTRACELLULAR PHOSPHORYLATION
: FILE REFERENCE: PR-0683 P
: CURRENT APPLICATION NUMBER: US/60/125,593
: CURRENT FILING DATE: 1999-03-18
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: PERL Program
: SEQ ID NO: 12
: LENGTH: 1573
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: 2415617
US-60-125-593-12

```

Qy	721	ctcgcgtctggcagacccttcgcgactctaccggggccgcactgacctctctctgtagcagaca	780
Db	721	ctcgcgtctggcagacccttcgcgactctaccggggccgcactgacctctctctgtagcagaca	780
Qy	781	gtggcggttcgcaaacacacgcgcctcccgccggtttgctgcacacagggccggctgtagatgcacac	840
Db	781	gtggcggttcgcaaacacacgcgcctcccgccggtttgctgcacacagggccggctgtagatgcacac	840
Qy	841	ctttgttcgcacatggagagaccacgagtcgacctctgcacatggagagtggtcccgagagagacattt	900
Db	841	ctttgttcgcacatggagagaccacgagtcgacctctgcacatggagagtggtcccgagagagacattt	900
Qy	901	gtgtctgcttcctggcagctgctgcactgcgacactggagagatgattttgtgcgaccacatg	960
Db	901	gtgtctgcttcctggcagctgctgcactgcgacactggagagatgattttgtgcgaccacatg	960
Qy	961	ggcgcgctgtgacgctggcgctcattgctctacgtgcggggggagagtgctcgtggc	1021
Db	961	ggcgcgctgtgacgctggcgctcattgctctacgtgcggggggagagtgctcgtggc	1021
Qy	1021	ggcgcgctgtgacgctggcgctcattgctctacgtgcggggggagagtgctcgtggc	1081
Db	1021	ggcgcgctgtgacgctggcgctcattgctctacgtgcggggggagagtgctcgtggc	1081

Query Match	Similarity	100.0%	Score 1573	DB 56	Length 1573
Best Local	Similarity	100.0%	Prod. No. 4,1e-303		
Matches 1573	Conservative	0	Mismatches	0	Indels
					Gaps
Qy	1	GGCCCAACAGCGGCGCTCTGACAGCCCGCTGCGAGCACACCGATAAGGACGCTGAAGCGACG	60		
Dy	1	GGCCCAACAGCGGCGCTCTGACAGCCCGCTGCGAGCACACCGATAAGGACGCTGAAGCGACG	60		
Qy	61	AGCGCGCGCCACAGGGCAGCGCCCAACAGCGCACAGGACCCCTGCGACAGCGGAGCGCGG	120		
Dy	61	AGCGCGCGCCACAGGGCAGCGCCCAACAGCGCACAGGACCCCTGCGACAGCGGAGCGCGG	120		
Qy	121	GTCGAGGTTATGATCCACAGCGGCGCGCGCGCGGCGTGTCTCCGCGGCGCTTCGCGCGT	180		
Dy	121	GTCGAGGTTATGATCCACAGCGGCGCGCGCGCGGCGTGTCTCCGCGGCGCTTCGCGCGT	180		
Qy	181	CTGAGTCTGCTGAACCCGCGCGCGCGGCGACAGGCGAAGGCGCTTCGACGCTTCGCGAGTAC	240		
Dy	181	CTGAGTCTGCTGAACCCGCGCGCGCGGCGACAGGCGAAGGCGCTTCGACGCTTCGCGAGTAC	240		
Qy	241	GTCGAGCGCCCTTTGGCTGAGGCTGAATCTCTTCACGCTGATGCTCACTGAGCGCGG	300		
Dy	241	GTCGAGCGCCCTTTGGCTGAGGCTGAATCTCTTCACGCTGATGCTCACTGAGCGCGG	300		

Qy	1022	TACTTCGTATATGTC	CCCCGTCGCTTC	CCGCTTGAGAG	CCCAAGAGTGGAAAGGTGTG	1144
Db	1021	ATTCCTGCTCGGCTCTTCTCTG	GCATGGAAGAGGACGATATGAGATGTGAATGCCCC	1086		
Qy	1081	TACTTCGTATATGTC	CCCCGTCGCTTC	CCGCTTGAGAG	CCCAAGAGTGGAAAGGTGTG	1144
Db	1081	TACTTCGTATATGTC	CCCCGTCGCTTC	CCGCTTGAGAG	CCCAAGAGTGGAAAGGTGTG	1144
Qy	1141	TTTTCAGTGTAGTGGAAATTGATGTTAG	CGAGGCGCTGACAGGGCCAGTGTCAACCCCAAC	1200		
Db	1141	TTTTCAGTGTAGTGGAAATTGATGTTAG	CGAGGCGCTGACAGGGCCAGTGTCAACCCCAAC	1200		
Qy	1201	TACTTCGTATGTC	TAGCGGTTGGTGGTGGAGACC	CGCCGACGTGGAAAGCCCCAGCAGATG	1266	
Db	1201	TACTTCGTATGTC	TAGCGGTTGGTGGTGGAGACC	CGCCGACGTGGAAAGCCCCAGCAGATG	1266	
Qy	1261	GCACGCGCCAGAAAGACCCCTTATGACC	CCCTGCGGCGCGGTGTGCTTATGTTACTTATGCA	1322		
Db	1261	GCACGCGCCAGAAAGACCCCTTATGACC	CCCTGCGGCGCGGTGTGCTTATGTTACTTATGCA	1322		
Qy	1321	GGACCGCTTCTCTCTCCCTAG	GGGTGAGAGGCGCTGTGCACAGCTCTCTGTGGGGGTGGAGG	1386		
Db	1321	GGACCGCTTCTCTCTCCCTAG	GGGTGAGAGGCGCTGTGCACAGCTCTCTGTGGGGGTGGAGG	1386		
Qy	1381	AGACCTCTCTGAGAGAAGGTGTGAGAAGGTGGAGGCTATATCTTTGGGGGGACAGGCCAGAAAT	1440			

22	1539	97.8	1772	24	US-09-629-669A-15393	Sequence 15393, A
23	1539	97.8	1715	17	US-10-111-729-1	Sequence 1, Appl1
24	1524.6	96.9	1797	40	US-09-338-667-876	Sequence 876, App
25	1524.6	96.9	1797	31	US-09-808-384-876	Sequence 876, App
26	1524.6	96.9	1797	31	US-09-813-206-806	Sequence 806, App
27	1515.4	96.3	1859	30	US-09-532-805-118	Sequence 118, App
28	1511.6	96.1	1600	20	US-09-784-810A-1	Sequence 1, Appl1
29	1479.2	94.0	1533	1	PCT-US01-05614-90	Sequence 90, Appl1
30	1479.2	94.0	1533	1	PCT-US98-114228-90	Sequence 90, Appl1
31	1479.2	94.0	1533	16	US-09-205-558-90	Sequence 90, Appl1
32	1479.2	94.0	1533	35	US-09-933-767-90	Sequence 90, Appl1
33	1479.2	94.0	1533	38	US-10-004-860-90	Sequence 90, Appl1
34	1479.2	94.0	1533	38	US-10-023-282-90	Sequence 90, Appl1
35	1430.2	90.9	1562	1	PCT-US01-16450-1180	Sequence 1180, App
36	1430.2	90.9	1562	1	PCT-US01-16450A-1180	Sequence 1180, App
37	1428.4	90.8	1438	18	US-09-488-725A-1539	Sequence 1539, App
38	1428.4	90.8	1438	25	US-09-653-550A-129	Sequence 129, App
39	1258.4	80.0	1272	1	PCT-US02-24623-1	Sequence 2, Appl1
40	1168.6	74.3	1205	36	US-09-959-897-1	Sequence 1, Appl1
41	1165	74.1	1173	80	US-06-364-517-1	Sequence 1, Appl1
42	1163.2	73.9	1200	18	US-09-471-775-5302	Sequence 5302, App
43	1163.2	73.9	1200	18	US-09-488-725A-5111	Sequence 5111, App
44	1155	73.4	1155	36	US-09-970-516-1	Sequence 1, Appl1
45	1155	73.4	1155	40	US-10-111-729-2	Sequence 2, Appl1

ALIGNMENTS

```

RESULT 1
US-09-937-060A-19
Sequence 19, Application US/09937060A
GENERAL INFORMATION:
APPLICANT: INCYTE PHARMACEUTICALS, INC.
APPLICANT: BANDMAN, Olga
APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: AZIMZAI, Valda
APPLICANT: LU, Dyung Alina M.
APPLICANT: AU-YOUNG, Janice
TITLE OF INVENTION: REGULATORS OF INTRACELLULAR PHOSPHORYLATION
FILE REFERENCE: PF-0683 PCT
CURRENT APPLICATION NUMBER: US/09/937,060A
CURRENT FILING DATE: 2002-04-15
PRIOR APPLICATION NUMBER: 60/125,593; 60/135,049; 60/143,188
PRIOR FILING DATE: 1999-03-18; 1999-05-20; 1999-07-09
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PERL Program
SEQ ID NO 19
LENGTH: 1573
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 2415617CB1
US-09-937-060A-19

Query Match      100.0%; Score 1573; DB 35; Length 1573;
Best Local Similarity 100.0%; Pred. No. 4,1e-303;
Matches 1573; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 GCCCAGAGCGGGCCCTCGAGCGCCGCGCTGGGACAGACCGAATAAGAGCTGAAGCGAG 60
Db 1 GCCCAGAGCGGGCCCTCGAGCAGCCGCCCTCGGGCAGCAGCGAATAAGAGCTGAAGGAGG 60
QY 61 AGCGCGCCGACAGGAGCGCGCCCGACAGCGCCAGGAGACCCCTTGCGCAGCGGAGCGCGG 120
Db 61 AGCGCGCCGACAGGAGCGCGCCCGACAGCGCCAGGAGACCCCTTGCGCAGCGGAGCGCGG 120
QY 121 GTGAGGTTATGCAATCCAGGAGGCGGCGCCCGGGGCGCTGCTCCCGGCGGCTGCGCGGTG 180
Db 121 GTGAGGTTATGCAATCCAGGAGGCGGCGCCCGGGGCGCTGCTCCCGGCGGCTGCGCGGTG 180

```

Db	121	GTGAGGTATGATCCAGCGGGGCGCCCCCGGGCGTGTCTCCCGGGCCCTGCCGCTG	180
Qy	181	CTGGTGTGTGAACCCGCGCGGCGCAAGGGCAAGGGCTTGGACCTTCTCGGAGTCAAC	240
Db	181	CTGGTGTGTGAACCCGCGCGGCGCGCAAGGGCAAGGGCTTGGACCTTCTCGGAGTCAAC	240
Qy	241	GTGAGGCCCCCTTTTGGCTGAGGCTGAATCTCCTTCACGCTGATGCTCACTGAGCGGGG	300
Db	241	GTGAGGCCCCCTTTTGGCTGAGGCTGAATCTCCTTCACGCTGATGCTCACTGAGCGGGG	300
Qy	301	AACACGCGGGGGGCGTGGGTGGGGAGACTGGGGCGGTGGGAGACTGTGGTGGTC	360
Db	301	AACACGCGGGGGGCGTGGGTGGGGAGACTGGGGCGGTGGGAGACTGTGGTGGTC	360
Qy	361	ATGCTGTGAGACGGGGCTGATGACGAGGATGGTAACGGGCTATGAGACGGGCTGACTGG	420
Db	361	ATGCTGTGAGACGGGGCTGATGACGAGGATGGTAACGGGCTATGAGACGGGCTGACTGG	420
Qy	421	GAGACCCGCATCCAGAAAGCCCTGTGTAGCTTCCACGACAGGCTCTGGCAACGCGTGCA	480
Db	421	GAGACCCGCATCCAGAAAGCCCTGTGTAGCTTCCACGACAGGCTCTGGCAACGCGTGCA	480
Qy	481	GCTTCCTGAAACATTAATGCTGGCTATGAGAGAGTACCATAAGAGACCTCGTAGCAAC	540
Db	481	GCTTCCTGAAACATTAATGCTGGCTATGAGAGAGTACCATAAGAGACCTCGTAGCAAC	540
Qy	541	TGCACGCTATTGCTGTGCCCGCGGCTGCTGCACCCATGACCTGCTCTGTGCACAG	600
Db	541	TGCACGCTATTGCTGTGCCCGCGGCTGCTGCACCCATGACCTGCTCTGTGCACAG	600
Qy	601	GCTTCGGGGCTGGCGCTCTTCTGTGTGCTGACCTGGGCTTCATTGTGATGTG	660
Db	601	GCTTCGGGGCTGGCGCTCTTCTGTGTGCTGACCTGGGCTTCATTGTGATGTG	660
Qy	661	GACCTAGAGAGTGAAGATGTGGGGCTGGGGGAGATGCGTTACATCTGGGCAACCTTC	720
Db	661	GACCTAGAGAGTGAAGATGTGGGGCTGGGGGAGATGCGTTACATCTGGGCAACCTTC	720
Qy	721	CTGCGTGTGGGACCCCTGGACCACTACCGCGGCGCACTGGCTTACCCTCCCTGTAGAGA	780
Db	721	CTGCGTGTGGGACCCCTGGACCACTACCGCGGCGCACTGGCTTACCCTCCCTGTAGAGA	780
Qy	781	GTGGGTTTCAAGACACTGCTCCCTCCGTTGTGTGCTACAGAGGGCCCGGTAGATGACAC	840
Db	781	GTGGGTTTCAAGACACTGCTCCCTCCGTTGTGTGCTACAGAGGGCCCGGTAGATGACAC	840
Qy	841	CTTGCGGCACGTGGAGGGAGGCAAGTGGCTCTGCATGTGACAGTGGTCCGACGAGACCTT	900
Db	841	CTTGCGGCACGTGGAGGGAGGCAAGTGGCTCTGCATGTGACAGTGGTCCGACGAGACCTT	900
Qy	901	GTCGTAGTCCGTGCACGTGCACACTGCGACCTTGGGCGATGAGATGTTTGTTCACCCATG	960
Db	901	GTCGTAGTCCGTGCACGTGCACACTGCGACCTTGGGCGATGAGATGTTTGTTCACCCATG	960
Qy	961	GGCCGCTGTGCAGCTGGCGTGCATGCAATCTGTTACGTGCGGGCGGAGTGTCTGTGCC	1020
Db	961	GGCCGCTGTGCAGCTGGCGTGCATGCAATCTGTTACGTGCGGGCGGAGTGTCTGTGCC	1020
Qy	1021	ATGCGTGGGCCCCCTTCTCGGGCCATGGAGAAAGGAGCCATATGGAGTATGAATGGCCC	1080
Db	1021	ATGCGTGGGCCCCCTTCTCGGGCCATGGAGAAAGGAGCCATATGGAGTATGAATGGCCC	1080
Qy	1081	TACTTGTGATATGTGCCCCGTGTGCGCTTCGCTTGGAGCCCAAGGATGGGAAAGTGTG	1140
Db	1081	TACTTGTGATATGTGCCCCGTGTGCGCTTCGCTTGGAGCCCAAGGATGGGAAAGTGTG	1140
Qy	1141	TTTTCAGTGTGATGGGGAATTGATGTTAGCGAGAGCCGTGCAGGGCCAGGTGCACCCAAC	1200
Db	1141	TTTTCAGTGTGATGGGGAATTGATGTTAGCGAGAGCCGTGCAGGGCCAGGTGCACCCAAC	1200
Qy	1201	TACTTGTGATGTGATGACGGGTGCGTGGAGACCCCGCCAGCTGGAGAGCCCGCAGATG	1260
Db	1201	TACTTGTGATGTGATGACGGGTGCGTGGAGACCCCGCCAGCTGGAGAGCCCGCAGATG	1260

Mon May.19 10:50:39 2003

us-09-937-060a-5.1rapn

Page 5

CURRENT APPLICATION NUMBER: US/60/452,680
CURRENT FILING DATE: 2003-03-07
NUMBER OF SEQ ID NOS: 116213
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12558
LENGTH: 595
TYPE: PRT
ORGANISM: Homo sapiens
US-60-452-680-12558

Query Match 44.0% Score 886.5; DB 7; Length 595;
Best Local Similarity 39.8%; Pred. No. 1,5e-74;

Matches 195; Conservative 61; Mismatches 107; Indels 127; Gaps 6;

QY 10 VLRPCRVYLVLNPRGCKKALQLEFRSHVOPPLAEAEISFTLMLTERNNHARELVRSSEEL 69
DB 117 LRPRLPRLLLVNPFGGGLAMOWCKNHVLPMTISEAGLSFNLIOERONHARELVQGLSL 176
QY 70 GRMDALVYVSGDGLMHEVYVNGIMERPDMETAIOKPLCSLPAGSGNALAASLNHYAGYEOV 129
DB 177 SEMDGLIVYSGDGLMHEVYVNGILDRPMEEAIVKMPYGIILPCSGNALAGAVNOHGFEPRA 236
QY 130 TNEDLTNGTLLCRLLSPMNLISLHTASGLRLEFSVLSLAWGFADVDLESEKRYRIGE 189
DB 237 LGDILLNCSLLLCGGGHPDLISLTVLASGSCFSLVAMGFSDVDIOSEFRALGS 296
QY 190 MREFTGLRLAALRTYGRGLAYLPVG-RVGSKTPA----- 224
DB 297 ARPTLGTVLGLATLHTYGRGLSYLPATVEPASPTPAHSILPRAKSELTLTPDPAPMAHSP 356
QY 225 ----- 224,
DB 357 LHRVSDDLPLPQALASPGSPERPLPLSLNGGPELAGDMGAGADAPLSPPDLLSSPP 416
QY 225 -----SPVYVQO-----GVDALHVLPLEEVPSPHMTYVP 253
DB 417 GSPKALHSFVSEGAFLVLPSSGLPLPTPDARVGAATCGPPDHLPLPGTLPDPDW-VTL 475
QY 254 DEDFVLVLAALSHLSEMFAPMRCAGVMHLFVYRAGVSRAMLRLFLAMEGRHME 313
DB 476 ECGFVLMALISPHIGADLVAAPIHARDODLVHLCVNRSGISRAALLRLFLAMEGRSHFS 535
QY 314 YECFVLYVYVVAFLERKDKGYFANDGELMYSVAVQGVIPNTFMVNSCVCEPPPSMK 373
DB 536 LCGFOLGYAAARAFRLERPLTRGVLTVDGEVEYGPLQAOMHPIGTLTGTG---PGCC- 590
QY 374 PQQMPPEEP 383
DB 591 -----PGREP 595

RESULT 10

US-60-453-135-7837
Sequence 7837, Application US/60453135

GENERAL INFORMATION:
APPLICANT: CARCILL, Michele

APPLICANT: TAKOBOVA, Olga
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

FILE REFERENCE: CLO01456
TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

CURRENT APPLICATION NUMBER: US/60/453,135
CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7837
LENGTH: 595

TYPE: PRT
ORGANISM: Homo sapiens

US-60-453-135-7837

Query Match 44.0% Score 886.5; DB 7; Length 595;
Best Local Similarity 39.8%; Pred. No. 1,5e-74;
Matches 195; Conservative 61; Mismatches 107; Indels 127; Gaps 6;

QY 10 VLRPCRVYLVLNPRGCKKALQLEFRSHVOPPLAEAEISFTLMLTERNNHARELVRSSEEL 69
DB 117 LRPRLPRLLLVNPFGGGLAMOWCKNHVLPMTISEAGLSFNLIOERONHARELVQGLSL 176
QY 70 GRMDALVYVSGDGLMHEVYVNGIMERPDMETAIOKPLCSLPAGSGNALAASLNHYAGYEOV 129
DB 177 SEMDGLIVYSGDGLMHEVYVNGILDRPMEEAIVKMPYGIILPCSGNALAGAVNOHGFEPRA 236
QY 130 TNEDLTNGTLLCRLLSPMNLISLHTASGLRLEFSVLSLAWGFADVDLESEKRYRIGE 189
DB 237 LGDILLNCSLLLCGGGHPDLISLTVLASGSCFSLVAMGFSDVDIOSEFRALGS 296
QY 190 MREFTGLRLAALRTYGRGLAYLPVG-RVGSKTPA----- 224
DB 297 ARPTLGTVLGLATLHTYGRGLSYLPATVEPASPTPAHSILPRAKSELTLTPDPAPMAHSP 356
QY 225 ----- 224
DB 357 LHRVSDDLPLPQALASPGSPERPLPLSLNGGPELAGDMGAGADAPLSPPDLLSSPP 416
QY 225 -----SPVYVQO-----GVDALHVLPLEEVPSPHMTYVP 253
DB 417 GSPKALHSFVSEGAFLVLPSSGLPLPTPDARVGAATCGPPDHLPLPGTLPDPDW-VTL 475
QY 254 DEDFVLVLAALSHLSEMFAPMRCAGVMHLFVYRAGVSRAMLRLFLAMEGRHME 313
DB 476 ECGFVLMALISPHIGADLVAAPIHARDODLVHLCVNRSGISRAALLRLFLAMEGRSHFS 535
QY 314 YECFVLYVYVVAFLERKDKGYFANDGELMYSVAVQGVIPNTFMVNSCVCEPPPSMK 373
DB 536 LCGFOLGYAAARAFRLERPLTRGVLTVDGEVEYGPLQAOMHPIGTLTGTG---PGCC- 590
QY 374 PQQMPPEEP 383
DB 591 -----PGREP 595

RESULT 11

US-60-453-050-7837
Sequence 7837, Application US/60453050

GENERAL INFORMATION:
APPLICANT: CARCILL, Michele

APPLICANT: LUKE, May
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

FILE REFERENCE: CLO01457
TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF

CURRENT APPLICATION NUMBER: US/60/453,050
CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7837
LENGTH: 595

TYPE: PRT
ORGANISM: Homo sapiens

US-60-453-050-7837

Query Match 44.0% Score 886.5; DB 7; Length 595;
Best Local Similarity 39.8%; Pred. No. 1,5e-74;
Matches 195; Conservative 61; Mismatches 107; Indels 127; Gaps 6;

QY 10 VLRPCRVYLVLNPRGCKKALQLEFRSHVOPPLAEAEISFTLMLTERNNHARELVRSSEEL 69
DB 117 LRPRLPRLLLVNPFGGGLAMOWCKNHVLPMTISEAGLSFNLIOERONHARELVQGLSL 176
QY 70 GRMDALVYVSGDGLMHEVYVNGIMERPDMETAIOKPLCSLPAGSGNALAASLNHYAGYEOV 129
DB 177 SEMDGLIVYSGDGLMHEVYVNGILDRPMEEAIVKMPYGIILPCSGNALAGAVNOHGFEPRA 236
QY 130 TNEDLTNGTLLCRLLSPMNLISLHTASGLRLEFSVLSLAWGFADVDLESEKRYRIGE 189
DB 237 LGDILLNCSLLLCGGGHPDLISLTVLASGSCFSLVAMGFSDVDIOSEFRALGS 296

PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 9811
LENGTH: 312
TYPE: PRT
ORGANISM: Human
US-09-949-016-9811

Query Match 81.4%; Score 1642; DB 5; Length 312;
Best Local Similarity 100.0%; Pred. No. 4,4e-146;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 DALVYMSGDGLMHEVYVNGLMERPMETAIQKPLCSLPAGSGNALAASLNHYAGYEQVTNE 132
DB 1 DALVYMSGDGLMHEVYVNGLMERPMETAIQKPLCSLPAGSGNALAASLNHYAGYEQVTNE 60
QY 133 DLTNCTLLCRRLSPMNLSTHTASGLRFLPSVLSLAWGFIADVDLSEKTRRIGEMRF 192
DB 61 DLTNCTLLCRRLSPMNLSTHTASGLRFLPSVLSLAWGFIADVDLSEKTRRIGEMRF 120
QY 193 TLGTFLRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPPVSHMTVV 252
DB 121 TLGTFLRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPPVSHMTVV 180
QY 253 PDEDVYVLAALSHLSGSEMFAPMGRCAAGVNHLYFVAGVSRAMLRLFLAMEKGRHM 312
DB 181 PDEDVYVLAALSHLSGSEMFAPMGRCAAGVNHLYFVAGVSRAMLRLFLAMEKGRHM 240
QY 313 EXECPYLVVPAVFAFRLEPKDKGVFAVDGELMVSFAVGOVHPNFMVSGCVPPEPSM 372
DB 241 EXECPYLVVPAVFAFRLEPKDKGVFAVDGELMVSFAVGOVHPNFMVSGCVPPEPSM 300
QY 373 KPQMPPEEPL 384
DB 301 KPQMPPEEPL 312

RESULT 7
US-60-453-135-13597
Sequence 13597, Application US/60453135
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: IAKOUBOVA, Olga
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
CURRENT APPLICATION NUMBER: US/60/453,135
CURRENT FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13597
LENGTH: 307
TYPE: PRT
ORGANISM: Homo sapiens
US-60-453-135-13597

Query Match 80.4%; Score 1620; DB 7; Length 307;
Best Local Similarity 100.0%; Pred. No. 5e-144;
Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 MSGDGLMHEVYVNGLMERPMETAIQKPLCSLPAGSGNALAASLNHYAGYEQVTNE 137
DB 1 MSGDGLMHEVYVNGLMERPMETAIQKPLCSLPAGSGNALAASLNHYAGYEQVTNE 60
QY 138 CTLLCRRLSPMNLSTHTASGLRFLPSVLSLAWGFIADVDLSEKTRRIGEMRF 197
DB 61 CTLLCRRLSPMNLSTHTASGLRFLPSVLSLAWGFIADVDLSEKTRRIGEMRF 120
QY 198 LRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPPVSHMTVV 257

DB 121 LRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPPVSHMTVV 180
QY 258 VLVALLSHLSGSEMFAPMGRCAAGVNHLYFVAGVSRAMLRLFLAMEKGRHMEYEC 317
DB 181 VLVALLSHLSGSEMFAPMGRCAAGVNHLYFVAGVSRAMLRLFLAMEKGRHMEYEC 240
QY 318 YLYVYVPAVFAFRLEPKDKGVFAVDGELMVSFAVGOVHPNFMVSGCVPPEPSM 377
DB 241 YLYVYVPAVFAFRLEPKDKGVFAVDGELMVSFAVGOVHPNFMVSGCVPPEPSM 300
QY 378 PPEEPL 384
DB 301 PPEEPL 307

RESULT 8
US-60-453-050-13597
Sequence 13597, Application US/60453050
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: LUKE, May
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF
CURRENT APPLICATION NUMBER: US/60/453,050
CURRENT FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13597
LENGTH: 307
TYPE: PRT
ORGANISM: Homo sapiens
US-60-453-050-13597

Query Match 80.4%; Score 1620; DB 7; Length 307;
Best Local Similarity 100.0%; Pred. No. 5e-144;
Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 MSGDGLMHEVYVNGLMERPMETAIQKPLCSLPAGSGNALAASLNHYAGYEQVTNE 137
DB 1 MSGDGLMHEVYVNGLMERPMETAIQKPLCSLPAGSGNALAASLNHYAGYEQVTNE 60
QY 138 CTLLCRRLSPMNLSTHTASGLRFLPSVLSLAWGFIADVDLSEKTRRIGEMRF 197
DB 61 CTLLCRRLSPMNLSTHTASGLRFLPSVLSLAWGFIADVDLSEKTRRIGEMRF 120
QY 198 LRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPPVSHMTVV 257
DB 121 LRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPPVSHMTVV 180
QY 258 VLVALLSHLSGSEMFAPMGRCAAGVNHLYFVAGVSRAMLRLFLAMEKGRHMEYEC 317
DB 181 VLVALLSHLSGSEMFAPMGRCAAGVNHLYFVAGVSRAMLRLFLAMEKGRHMEYEC 240
QY 318 YLYVYVPAVFAFRLEPKDKGVFAVDGELMVSFAVGOVHPNFMVSGCVPPEPSM 377
DB 241 YLYVYVPAVFAFRLEPKDKGVFAVDGELMVSFAVGOVHPNFMVSGCVPPEPSM 300
QY 378 PPEEPL 384
DB 301 PPEEPL 307

RESULT 9
US-60-452-680-12558
Sequence 12558, Application US/60452680
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: GRUBE, Andrew
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF

Mon May. 19 10:50:39 2003

us-09-937-060a-5.i.rapn

Page 3

Best Local Similarity 99.2%; Pred. No. 2, 1e-180;
Matches 361; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
OY 1 MPDAGGPRVLPKPRVLYLNLPRGCKKALQFRSHVOPLLAEAEISFTLMLERRNHA 60
    |||
Db 1 MDPAAGGRVLPKPRVLYLNLPRGCKKALQFRSHVOPLLAEAEISFTLMLERRNHA 60
OY 61 RELVSEELGRMDALVWSDGLMHEVNGIMERPDMETAIOKPLCSLPASSGNALAST 120
    |||
Db 61 RELVSEELGRMDALVWSDGLMHEVNGIMERPDMETAIOKPLCSLPASSGNALAST 120
OY 121 NHVAGEQVNTNEDLLTNCJLLCRLLSPNNLLSLHTASGLRFLSVLSLWGTADVDLE 180
    |||
Db 121 NHVAGEQVNTNEDLLTNCJLLCRLLSPNNLLSLHTASGLRFLSVLSLWGTADVDLE 180
OY 181 SEKVRRLGEMRFTLGTFRLAALRTYGRRLAYLPVGRVSKTPASPVVVOGPDADHLVP 240
    |||
Db 181 SEKVRRLGEMRFTLGTFRLAALRTYGRRLAYLPVGRVSKTPASPVVVOGPDADHLVP 240
OY 241 LEEPVSHMTVPDDEFLVLLALSHLSEMFAPMGRCAGVMHLFVYRAGVSRAML 300
    |||
Db 241 LEEPVSHMTVPDDEFLVLLALSHLSEMFAPMGRCAGVMHLFVYRAGVSRAML 300
OY 301 RLFLAMEKGRHMEYECPTLYVVPVAVARLEPKDGKGFVAVDGLMVSAAVVOGHPNTEFW 360
    |||
Db 301 RLFLAMEKGRHMEYECPTLYVVPVAVARLEPKDGKGFVAVDGLMVSAAVVOGHPNTEFW 360
OY 361 MMSGVEPPPSMKPOQMPPEEPL 384
    |||
Db 361 MMSGVEPPPSMKPOQMPPEEPL 384
```

RESULT 4

```
US-10-348-052-21
; Sequence 21, Application US/10348052
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION
; FILE REFERENCE: 200116.405
; CURRENT APPLICATION NUMBER: US/10/348.052
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-348-052-21
```

Query Match 95.4%; Score 1923; DB 6; Length 368;

Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 17 VLVLLNPRGCKKALQFRSHVOPLLAEAEISFTLMLERRNHAELVRSSEELGRMDALV 76
    |||
Db 1 VLVLLNPRGCKKALQFRSHVOPLLAEAEISFTLMLERRNHAELVRSSEELGRMDALV 76
OY 77 VMSGDGLMHEVNGIMERPDMETAIOKPLCSLPASSGNALASTNHVAGEQVNTNEDLLT 136
    |||
Db 61 VMSGDGLMHEVNGIMERPDMETAIOKPLCSLPASSGNALASTNHVAGEQVNTNEDLLT 120
OY 137 NCTLLLCRRLLSPNNLLSLHTASGLRFLSVLSLWGTADVDLESEKVRRLGEMRFTLG 196
    |||
Db 121 NCTLLLCRRLLSPNNLLSLHTASGLRFLSVLSLWGTADVDLESEKVRRLGEMRFTLG 180
OY 197 FLRLAALRTYGRRLAYLPVGRVSKTPASPVVVOGPDADHLVPLEEPVSHMTVPDDE 256
    |||
Db 181 FLRLAALRTYGRRLAYLPVGRVSKTPASPVVVOGPDADHLVPLEEPVSHMTVPDDE 240
OY 257 FVLVLLALSHLSEMFAPMGRCAGVMHLFVYRAGVSRAMLRLFLAMEKGRHMEYEC 316
    |||
```

```
Db 241 FVLVLLALSHLSEMFAPMGRCAGVMHLFVYRAGVSRAMLRLFLAMEKGRHMEYEC 300
OY 317 PVLVYVPVAVARLEPKDGKGFVAVDGLMVSAAVVOGHPNTEFWVMSGVEPPPSMKPOQ 376
    |||
Db 301 PVLVYVPVAVARLEPKDGKGFVAVDGLMVSAAVVOGHPNTEFWVMSGVEPPPSMKPOQ 360
OY 377 MPEPEEPL 384
    |||
Db 361 MPEPEEPL 368
```

RESULT 5

```
US-10-264-237-2585
; Sequence 2585, Application US/10264237
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4331
; CURRENT APPLICATION NUMBER: US/10/264.237
; PRIOR FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/0501/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: Patentin Ver. 3.1
; SEQ ID NO: 2585
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-2585
```

Query Match 86.9%; Score 1752; DB 6; Length 333;

Best Local Similarity 100.0%; Pred. No. 2, 1e-156;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 52 MLTERRNHARELVSEELGRMDALVWSDGLMHEVNGIMERPDMETAIOKPLCSLPAG 111
    |||
Db 1 MLTERRNHARELVSEELGRMDALVWSDGLMHEVNGIMERPDMETAIOKPLCSLPAG 60
OY 112 SGNALAASLNHYAGYBOVTNEDLLTNCJLLCRLLSPNNLLSLHTASGLRFLSVLSLW 171
    |||
Db 61 SGNALAASLNHYAGYBOVTNEDLLTNCJLLCRLLSPNNLLSLHTASGLRFLSVLSLW 120
OY 172 GFADVDLESEKVRRLGEMRFTLGTFRLAALRTYGRRLAYLPVGRVSKTPASPVVVOO 231
    |||
Db 121 GFADVDLESEKVRRLGEMRFTLGTFRLAALRTYGRRLAYLPVGRVSKTPASPVVVOO 180
OY 232 GPVDALVPLEEPVSHMTVPDDEFLVLLALSHLSEMFAPMGRCAGVMHLFVYR 291
    |||
Db 181 GPVDALVPLEEPVSHMTVPDDEFLVLLALSHLSEMFAPMGRCAGVMHLFVYR 240
OY 292 ACVSRAMLRLFLAMEKGRHMEYECPTLYVVPVAVARLEPKDGKGFVAVDGLMVSAAV 351
    |||
Db 241 ACVSRAMLRLFLAMEKGRHMEYECPTLYVVPVAVARLEPKDGKGFVAVDGLMVSAAV 300
OY 352 GOVHNRYFMVMSGVEPPPSMKPOQMPPEEPL 384
    |||
Db 301 GOVHNRYFMVMSGVEPPPSMKPOQMPPEEPL 333
```

RESULT 6

```
US-09-949-016-9811
; Sequence 9811, Application US/09949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949.016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

ORGANISM: Homo sapiens
PCT-US03-02588-26

Query Match 100.0%; Score 2016; DB 1; Length 384;
Best Local Similarity 100.0%; Pred. No. 3.7e-181;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGPRGVLPKPCRYLVLLNPRGKGAALQLFPSHVQPLAEAEISFTMLTERRNA 60
DB 1 MDPAGGPRGVLPKPCRYLVLLNPRGKGAALQLFPSHVQPLAEAEISFTMLTERRNA 60

QY 61 RELVSEELGRMDALVWMSGDLMEHVNGLMERPDMETAIQKPLCSIPASGNALASL 120
DB 61 RELVSEELGRMDALVWMSGDLMEHVNGLMERPDMETAIQKPLCSIPASGNALASL 120

QY 121 NHYAGEQVNTMEDLLTNTLLCRRLSPNNLSLHTASGRLFSVLSLAWGFIADVLE 180
DB 121 NHYAGEQVNTMEDLLTNTLLCRRLSPNNLSLHTASGRLFSVLSLAWGFIADVLE 180

QY 181 SEKYRRLGEMRFTLTGFLRLAALRTYGRRLAYLPYGRVSKTPASPVVVOGSPVDAHLP 240
DB 181 SEKYRRLGEMRFTLTGFLRLAALRTYGRRLAYLPYGRVSKTPASPVVVOGSPVDAHLP 240

QY 241 LEEPVSHMTVVPDEDVFLVLLALSHLSEMFAPMGRCAAGVNHLEFYRAGVSRAML 300
DB 241 LEEPVSHMTVVPDEDVFLVLLALSHLSEMFAPMGRCAAGVNHLEFYRAGVSRAML 300

QY 301 RLFLAMEGRHMEYECPLVYVVAFRLEPKDGGVAVDGLMVSBAVGOVHPNFW 360
DB 301 RLFLAMEGRHMEYECPLVYVVAFRLEPKDGGVAVDGLMVSBAVGOVHPNFW 360

QY 361 MVSGCPEPPPSMKPOOMPPEEPL 384
DB 361 MVSGCPEPPPSMKPOOMPPEEPL 384

RESULT 2
US-10-354-358-26
Sequence 26, Application US/10354358
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc
APPLICANT: Hunter, John Joseph
APPLICANT: Macbeth, Kyle J.
APPLICANT: Tsai, Fong-Ying
APPLICANT: Lesoon, Andrea
APPLICANT: Lightcap, Eric S.
APPLICANT: Williamson, Mark
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
TITLE OF INVENTION: 8990, 2100, 9288, 64968, 10480, 20893, 33230, 1586, 9943,
TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6179, 21225, 1420, 32236, 2099,
FILE REFERENCE: MP102-020P1NOMNIM
CURRENT APPLICATION NUMBER: US/10-354,358
CURRENT FILING DATE: 2003-01-30
PRIOR APPLICATION NUMBER: US 60/353,600
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: US 60/364,517
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/371,075
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: US 60/371,507
PRIOR FILING DATE: 2002-04-10
PRIOR APPLICATION NUMBER: US 60/372,984
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/374,194
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/382,995
PRIOR FILING DATE: 2002-05-24

PRIOR APPLICATION NUMBER: US 60/385,023
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/388,853
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 60/389,395
PRIOR FILING DATE: 2002-06-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 26
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
US-10-354-358-26

Query Match 100.0%; Score 2016; DB 6; Length 384;
Best Local Similarity 100.0%; Pred. No. 3.7e-181;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGPRGVLPKPCRYLVLLNPRGKGAALQLFPSHVQPLAEAEISFTMLTERRNA 60
DB 1 MDPAGGPRGVLPKPCRYLVLLNPRGKGAALQLFPSHVQPLAEAEISFTMLTERRNA 60

QY 61 RELVSEELGRMDALVWMSGDLMEHVNGLMERPDMETAIQKPLCSIPASGNALASL 120
DB 61 RELVSEELGRMDALVWMSGDLMEHVNGLMERPDMETAIQKPLCSIPASGNALASL 120

QY 121 NHYAGEQVNTMEDLLTNTLLCRRLSPNNLSLHTASGRLFSVLSLAWGFIADVLE 180
DB 121 NHYAGEQVNTMEDLLTNTLLCRRLSPNNLSLHTASGRLFSVLSLAWGFIADVLE 180

QY 181 SEKYRRLGEMRFTLTGFLRLAALRTYGRRLAYLPYGRVSKTPASPVVVOGSPVDAHLP 240
DB 181 SEKYRRLGEMRFTLTGFLRLAALRTYGRRLAYLPYGRVSKTPASPVVVOGSPVDAHLP 240

QY 241 LEEPVSHMTVVPDEDVFLVLLALSHLSEMFAPMGRCAAGVNHLEFYRAGVSRAML 300
DB 241 LEEPVSHMTVVPDEDVFLVLLALSHLSEMFAPMGRCAAGVNHLEFYRAGVSRAML 300

QY 301 RLFLAMEGRHMEYECPLVYVVAFRLEPKDGGVAVDGLMVSBAVGOVHPNFW 360
DB 301 RLFLAMEGRHMEYECPLVYVVAFRLEPKDGGVAVDGLMVSBAVGOVHPNFW 360

QY 361 MVSGCPEPPPSMKPOOMPPEEPL 384
DB 361 MVSGCPEPPPSMKPOOMPPEEPL 384

RESULT 3
US-09-949-016-7026
Sequence 7026, Application US/09949016
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09-949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7026
LENGTH: 384
TYPE: PRT
ORGANISM: Human
US-09-949-016-7026

Query Match 99.6%; Score 2008; DB 5; Length 384;


```

Db 61 VMSGDGLMEVYVNGLMERPDWETAIOKPLCSLPGSGNALASLHNYAGYEOVTNEDLLT 120
QY 137 NCTLLCRRLSPMNLSTLTASGLRFLFSVLSLAWFIADVDLSEKRYRIGEMRFLGT 136
Db 121 NCTLLCRRLSPMNLSTLTASGLRFLFSVLSLAWFIADVDLSEKRYRIGEMRFLGT 180
QY 197 FLRLAALRTYRGRLAYLPVGVSGSKTPASPVYVQOCPVDAHLVPLEEVPVSHWTVPDED 256
Db 181 FLRLAALRTYRGRLAYLPVGVSGSKTPASPVYVQOCPVDAHLVPLEEVPVSHWTVPDED 240
QY 257 FVLVYVALLSHLSGEMFAPMGRCAGVNHLEFVYRAGVSRAHLFLFLAMEKGRHMETEC 316
Db 241 FVLVYVALLSHLSGEMFAPMGRCAGVNHLEFVYRAGVSRAHLFLFLAMEKGRHMETEC 300
QY 317 FVLVYVYVVAERFLERPKDGKGFVAVDGLMVSSEAVVQGVHPNTFMVNSGCEVPPPSMKPOO 376
Db 301 FVLVYVYVVAERFLERPKDGKGFVAVDGLMVSSEAVVQGVHPNTFMVNSGCEVPPPSMKPOO 360
QY 377 MPEPEPL 384
Db 361 MPEPEPL 368

```

```

RESULT 14
PCT-US01-16450-2585
: Sequence 2585, Application PC/TUS0116450
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA131PCT
: CURRENT APPLICATION NUMBER: PCT/US01/16450
: CURRENT FILING DATE: 2001-05-01
: PRIOR APPLICATION NUMBER: 60/205,515
: NUMBER OF SEQ ID NOS: 2820
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 2585
: LENGTH: 333
: TYPE: PRT
: ORGANISM: Homo sapiens
PCT-US01-16450-2585

```

```

Query Match      86.9%; Score 1752; DB 1; Length 333;
Best Local Similarity 100.0%; Pred. No. 8.7e-176;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 52 MLTERRNHARELVSEELGRMDALVYMSGDGLMHEVYVNGLMERPDWETAIOKPLCSLPAG 111
Db 1 MLTERRNHARELVSEELGRMDALVYMSGDGLMHEVYVNGLMERPDWETAIOKPLCSLPAG 60
QY 112 SGNALASLNHYAGYEOVTNEDLLTNCITLLCRRLSPMNLSTLTASGLRFLFSVLSLAW 171
Db 61 SGNALASLNHYAGYEOVTNEDLLTNCITLLCRRLSPMNLSTLTASGLRFLFSVLSLAW 120
QY 172 GFLADVDESEKRYRIGEMRFLGTFLRLAALRTYRGRLAYLPVGVSGSKTPASPVYVQO 231
Db 121 GFLADVDESEKRYRIGEMRFLGTFLRLAALRTYRGRLAYLPVGVSGSKTPASPVYVQO 180
QY 232 GPVDAHLVPLEEVPVSHWTVPDEDVYVYVALLSHLSGEMFAPMGRCAGVNHLEFYR 291
Db 181 GPVDAHLVPLEEVPVSHWTVPDEDVYVYVALLSHLSGEMFAPMGRCAGVNHLEFYR 240
QY 292 AGVSRAMLRLFLAMEKGRHMETECRYLVYVVAERFLERPKDGKGFVAVDGLMVSSEAVVQ 351
Db 241 AGVSRAMLRLFLAMEKGRHMETECRYLVYVVAERFLERPKDGKGFVAVDGLMVSSEAVVQ 300
QY 352 GOVHPNTFMVNSGCEVPPPSMKPOOMPPEPEPL 384
Db 301 GOVHPNTFMVNSGCEVPPPSMKPOOMPPEPEPL 333

```

RESULT 15

```

PCT-US01-16450A-2585
: Sequence 2585, Application PC/TUS0116450A
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA131PCT
: CURRENT APPLICATION NUMBER: PCT/US01/16450A
: CURRENT FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: 60/205,515
: NUMBER OF SEQ ID NOS: 2820
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 2585
: LENGTH: 333
: TYPE: PRT
: ORGANISM: Homo sapiens
PCT-US01-16450A-2585

```

```

Query Match      86.9%; Score 1752; DB 1; Length 333;
Best Local Similarity 100.0%; Pred. No. 8.7e-176;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 52 MLTERRNHARELVSEELGRMDALVYMSGDGLMHEVYVNGLMERPDWETAIOKPLCSLPAG 111
Db 1 MLTERRNHARELVSEELGRMDALVYMSGDGLMHEVYVNGLMERPDWETAIOKPLCSLPAG 60
QY 112 SGNALASLNHYAGYEOVTNEDLLTNCITLLCRRLSPMNLSTLTASGLRFLFSVLSLAW 171
Db 61 SGNALASLNHYAGYEOVTNEDLLTNCITLLCRRLSPMNLSTLTASGLRFLFSVLSLAW 120
QY 172 GFLADVDESEKRYRIGEMRFLGTFLRLAALRTYRGRLAYLPVGVSGSKTPASPVYVQO 231
Db 121 GFLADVDESEKRYRIGEMRFLGTFLRLAALRTYRGRLAYLPVGVSGSKTPASPVYVQO 180
QY 232 GPVDAHLVPLEEVPVSHWTVPDEDVYVYVALLSHLSGEMFAPMGRCAGVNHLEFYR 291
Db 181 GPVDAHLVPLEEVPVSHWTVPDEDVYVYVALLSHLSGEMFAPMGRCAGVNHLEFYR 240
QY 292 AGVSRAMLRLFLAMEKGRHMETECRYLVYVVAERFLERPKDGKGFVAVDGLMVSSEAVVQ 351
Db 241 AGVSRAMLRLFLAMEKGRHMETECRYLVYVVAERFLERPKDGKGFVAVDGLMVSSEAVVQ 300
QY 352 GOVHPNTFMVNSGCEVPPPSMKPOOMPPEPEPL 384
Db 301 GOVHPNTFMVNSGCEVPPPSMKPOOMPPEPEPL 333

```

Search completed: May 9, 2003, 16:52:50
Job time : 152 secs

NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin version 3.1
SEQ ID NO 3
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1)..(384)
OTHER INFORMATION: SEQ ID NO 3 is the peptide sequence of hSPHK1 in Fig. 3, correspo
OTHER INFORMATION: nding to amino acid residue 1 to 384 of Homo sapiens SPHK-1 of Ge
OTHER INFORMATION: nbank sequence Accession Number AAF73423.
PUBLICATION INFORMATION:
AUTHORS: Nava et al.
TITLE: Functional characterization of human spingosine kinase-1.
JOURNAL: FEBS Lett.
VOLUME: 473
ISSUE: 1
PAGES: 81-84
DATE: 2000
DATABASE ACCESSION NUMBER: AAF73423
DATABASE ENTRY DATE: 2000-06-01
RELEVANT RESIDUES: (1)..(384)
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: AAF73423
DATABASE ENTRY DATE: 2000-06-01
RELEVANT RESIDUES: (1)..(384)
US-09-796-487-3

Query Match 99.6%; Score 2008; DB 21; Length 384;
Best Local Similarity 99.2%; Pred. No. 8.5e-203;
Matches 381; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDPAGPRGVLPKPCRVYLVLNPRGKGKALQIFRSHVQPLAEAFISFTMLTERRNA 60
DB 1 MDPAGPRGVLPKPCRVYLVLNPRGKGKALQIFRSHVQPLAEAFISFTMLTERRNA 60
QY 61 RELVRSSEELGRMDALVVMGSDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALAASI 120
DB 61 RELVRSSEELGRMDALVVMGSDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALAASI 120
QY 121 NHYAGEQVTNEDLTNCTLLCRLLSPMNLISHTASGLRFSVLSLWGFIAVDLE 180
DB 121 NHYAGEQVTNEDLTNCTLLCRLLSPMNLISHTASGLRFSVLSLWGFIAVDLE 180
QY 181 SEKRYRLGEMRFTLGFRLAALRTYRGLATLPVGRVSKTPASPVVVOOGVDAHLVP 240
DB 181 SEKRYRLGEMRFTLGFRLAALRTYRGLATLPVGRVSKTPASPVVVOOGVDAHLVP 240
QY 241 LEEPVPSHMTVPDEDFVYLALHSHLSGSEMFAPMGCAAGVHMLFYVRAGVSRAML 300
DB 241 LEEPVPSHMTVPDEDFVYLALHSHLSGSEMFAPMGCAAGVHMLFYVRAGVSRAML 300
QY 301 RLFLAMEKGRHMEYECPLYVYVVAFLRLEPKDGKGFVAVDGLWVSEAVOGVHPNYTM 360
DB 301 RLFLAMEKGRHMEYECPLYVYVVAFLRLEPKDGKGFVAVDGLWVSEAVOGVHPNYTM 360
QY 361 MVSQGVPEPPPSMKPQOMPPEEPL 384
DB 361 MVSQGVPEPPPSMKPQOMPPEEPL 384

RESULT 12
US-09-784-810A-2

Sequence 2, Application US/09784810A
GENERAL INFORMATION:
APPLICANT: RASTELL, LUCA
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
TITLE OF INVENTION: SAME
FILE REFERENCE: 10716-08
CURRENT APPLICATION NUMBER: US/09/784,810A
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/182,360

PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 60/191,261
PRIOR FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
US-09-784-810A-2

Query Match 97.3%; Score 1962; DB 21; Length 384;
Best Local Similarity 97.7%; Pred. No. 6.3e-198;
Matches 375; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 MDPAGPRGVLPKPCRVYLVLNPRGKGKALQIFRSHVQPLAEAFISFTMLTERRNA 60
DB 1 MDPAGPRGVLPKPCRVYLVLNPRGKGKALQIFRSHVQPLAEAFISFTMLTERRNA 60
QY 61 RELVRSSEELGRMDALVVMGSDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALAASI 120
DB 61 RELVRSSEELGRMDALVVMGSDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALAASI 120
QY 121 NHYAGEQVTNEDLTNCTLLCRLLSPMNLISHTASGLRFSVLSLWGFIAVDLE 180
DB 121 NHYAGEQVTNEDLTNCTLLCRRLSPMNLISHTASGLRFSVLSLWGFIAVDLE 180
QY 181 SEKRYRLGEMRFTLGFRLAALRTYRGLATLPVGRVSKTPASPVVVOOGVDAHLVP 240
DB 181 SEKRYRLGEMRFTLGFRLAALRTYRGLATLPVGRVSKTPASPVVVOOGVDAHLVP 240
QY 241 LEEPVPSHMTVPDEDFVYLALHSHLSGSEMFAPMGCAAGVHMLFYVRAGVSRAML 300
DB 241 LEEQVSHMQVYVDEDFVYLALHSHLSGSEMFAPMGCAAGVHMLFYVRAGVSRAML 300
QY 301 RLFLAMEKGRHMEYECPLYVYVVAFLRLEPKDGKGFVAVDGLWVSEAVOGVHPNYTM 360
DB 301 RLFLAMEKGRHMEYECPLYVYVVAFLRLEPKDGKGFVAVDGLWVSEAVOGVHPNYTM 360
QY 361 MVSQGVPEPPPSMKPQOMPPEEPL 384
DB 361 MVSQGVPEPPPSMKPQOMPPEEPL 384

RESULT 13
US-10-053-510-21

Sequence 21, Application US/10053510
GENERAL INFORMATION:
APPLICANT: Saba, Julie D.
TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
TITLE OF INVENTION: METHODS OF USE THEREFOR
FILE REFERENCE: 200116.402C2
CURRENT APPLICATION NUMBER: US/10/053,510
CURRENT FILING DATE: 2002-01-17
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 368
TYPE: PRT
ORGANISM: Homo sapiens
US-10-053-510-21

Query Match 95.4%; Score 1923; DB 24; Length 368;
Best Local Similarity 100.0%; Pred. No. 8e-194;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 VLVLNPRGKGKALQIFRSHVQPLAEAFISFTMLTERNHARLVSEELGRMDALV 76
DB 1 VLVLNPRGKGKALQIFRSHVQPLAEAFISFTMLTERNHARLVSEELGRMDALV 60
QY 77 VMSGDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALAASLNHYAGEQVTNEDLT 136

Mon May 19 10:50:39 2003

us-09-937-060a-5. rapm

Page 5

Qy	241	LEEPVSHMTVYVDEDEFFVYLALHSHLSEMPAAMPKRCACAGMHLFFVRAGVSRAALL	3000
Db	241	LEEPVSHMTVYVDEDEFFVYLALHSHLSEMPAAMPKRCACAGMHLFFVRAGVSRAALL	3000
Qy	301	RIEFLAMEGRHMEYECPPYLVYVPAVARLEPPKKGAGVAVNGELAMSEAVQCOVHPNFW	360
Db	301	RIEFLAMEGRHMEYECPPYLVYVPAVARLEPPKKGAGVAVNGELAMSEAVQCOVHPNFW	360
Qy	361	MVSGGVPPPSNMKROOMPPPEEPL	384
Db	361	MVSGGVPPPSNMKROOMPPPEEPL	384

```

RESULT 9
US-09-488-725A-3325
? Sequence 3325, Application US/09488725A
? GENERAL INFORMATION:
?   APPLICANT: Hyseq Inc
?   TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
?   FILE REFERENCE: 784FLPCT
?   CURRENT APPLICATION NUMBER: US/09/488,725A
?   CURRENT FILING DATE: 2000-12-22
?   PRIOR APPLICATION NUMBER: US/09/488,725
?   PRIOR FILING DATE: 2000-01-21
?   PRIOR APPLICATION NUMBER: US09/552,317
?   PRIOR FILING DATE: 2000-04-25
?   PRIOR APPLICATION NUMBER: US09/598,042
?   PRIOR FILING DATE: 2000-06-20
?   PRIOR APPLICATION NUMBER: US09/620,312
?   PRIOR FILING DATE: 2000-07-19
?   PRIOR APPLICATION NUMBER: US09/653,450
?   PRIOR FILING DATE: 2000-08-31
?   PRIOR APPLICATION NUMBER: US09/662,191
?   PRIOR FILING DATE: 2000-09-14
?   PRIOR APPLICATION NUMBER: US09/693,036
?   PRIOR FILING DATE: 2000-10-19
?   PRIOR APPLICATION NUMBER: US09/727,344
?   PRIOR FILING DATE: 2000-11-29
?   NUMBER OF SEQ. ID NOS: 7144
?   SOFTWARE: pt-fl_genes_b Versions 1.0
?   SEQ. ID NO. 3325
?   LENGTH: 384
?   TYPE: PRT
?   ORGANISM: Homo sapiens
US-09-488-725A-3325

```

Query match	99.88;	Score 2011;	DB 18;	Length 384;
-------------	--------	-------------	--------	-------------

Best Local Similarity 99.7%; Pred.No. 4.1e-203;
Matches 383; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MACCGPCGVLP	PPRCVIVL	ILNPPGGKALQI	FRSHVQ	PLLAFAEISFLIM	TERRNA	60		
Db	1	MDPAGCPGVLP	PRCHVIVL	ILNPPGGKALQI	FRSHVQ	PLLAFAEISFLIM	TERRNA	60		
Qy	61	RELVAESELGHWALVYMSGDS	LHEVYNV	ILMPRMETAL	IQKPLCS	PRSGSNLA	SL	120		
Db	61	RELVAESELGHWALVYMSGDS	LHEVYNV	ILMPRMETAL	IQKPLCS	PRSGSNLA	SL	120		
Qy	121	NHAYAEQVYNEDL	INCLNCL	LRLLSPMLLS	SLHTASGL	IRSVLS	SLAMFINDVLE	180		
Db	121	NHAYAEQVYNEDL	INCLNCL	LRLLSPMLLS	SLHTASGL	IRSVLS	SLAMFINDVLE	180		
Qy	181	SEKRYRRLGEMAF	LTGFL	LMALAL	RTYGRAL	Y	PLVGRNGSKTPAS	SPVYVQOGVNAHLVP	240	
Db	181	SEKRYRRLGEMAF	LTGFL	LMALAL	RTYGRAL	Y	PLVGRNGSKTPAS	SPVYVQOGVNAHLVP	240	
Qy	241	LEEPVPSHMTV	VPDEDF	VYLLAL	LSHSDGEM	FAAPKGRCA	AGYMHLE	YFRAGVSRA	300	
Db	241	LEEPVPSHMTV	VPDEDF	VYLLAL	LSHSDGEM	FAAPKGRCA	AGYMHLE	YFRAGVSRA	300	
Qy	301	RLFLAMEGRHME	ECQPV	LVYVPVA	FLRLEPKDKG	QVAPVQDGL	MSA	SEAYQGVHPNT	FW	360
Db	301	RLFLAMEGRHME	ECQPV	LVYVPVA	FLRLEPKDKG	QVAPVQDGL	MSA	SEAYQGVHPNT	FW	360

QY	361	MVSGCVEPPPSWKPDQMPPEEPL	384
Db	361	MVSGCVEPPPSWKPDQMPPEEPL	384

```

RESULT 10
PCT-US02-24623-10
Sequence 10, Application PC/TUS0224623
GENERAL INFORMATION:
APPLICANT: EXELIXIS, INC.
TITLE OF INVENTION: SPKRS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
FILE REFERENCE: EX02-104
CURRENT APPLICATION NUMBER: PCT/US02/24623
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/310,362
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: US 60/357,501
PRIOR FILING DATE: 2002-02-15
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-24623-10

```

Query Match	99.68;	Score 2008;	DB 1;	Length 384;
-------------	--------	-------------	-------	-------------

Best Local Similarity	99.28	Pred. No.	8.5e-203
Matches	381	Conservative	2
		Mismatches	1
		Indels	0
		Gaps	0

QY	1	MDPAGGRRGVLPKPCRVLYLVLPNPGCKGKALDLPFSHYQPLLAFAEISTMLUTERRNHA	60
QY	1	MDPAGGRRGVLPKPCRVLYLVLPNPGCKGKALDLPFSHYQPLLAFAEISTMLUTERRNHA	60
Db	1	MDPAGGRRGVLPKPCRVLYLVLPNPGCKGKALDLPFSHYQPLLAFAEISTMLUTERRNHA	60
QY	61	RELVRSEELCRDALVWMSGDGLMEHVNGMLHEWRDPMTALOKTCLSPAGSGMLAASL	120
Db	61	RELVRSEELCRDALVWMSGDGLMEHVNGMLHEWRDPMTALOKTCLSPAGSGMLAASL	120
QY	121	NHYAGYEQVNTNEDLLNCTLLLCRRLLSPNNLLSHTASGLRFLFSVLSTAMGFINDVLE	180
Db	121	NHYAGYEQVNTNEDLLNCTLLLCRRLLSPNNLLSHTASGLRFLFSVLSTAMGFINDVLE	180
QY	181	SEKRYRLGKRFPLTGLRLALRTYRGRLAYLPVGRVGSSTPASPVYVQGGPVDAILVP	240
Db	181	SEKRYRLGKRFPLTGLRLALRTYRGRLAYLPVGRVGSSTPASPVYVQGGPVDAILVP	240
QY	241	LEEPVPSHMTVPVDEDFVLTALLSHSLGSEMFAPMGRCAAGYMHLPYVRAGVSRAML	300
Db	241	LEEPVPSHMTVPVDEDFVLTALLSHSLGSEMFAPMGRCAAGYMHLPYVRAGVSRAML	300
QY	301	RLFLTAMKGGHMEYECPLVYVPVAFRLRPPKDKGKGVAVDGEMLTSEAVOGVHPNPFW	360
Db	301	RLFLTAMKGGHMEYECPLVYVPVAFRLRPPKDKGKGVAVDGEMLTSEAVOGVHPNPFW	360
QY	361	MVSGCVPPEPSMKPDOMPPEEPL	384
Db	361	MVSGCVPPEPSMKPDOMPPEEPL	384

RESULT 11
 US-09-796-487-3
 Sequence 3, Application US/09796487
 GENERAL INFORMATION:
 Applicant: Sologel, Sarah
 TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
 FILE REFERENCE: 07320001a (2033957-0001)
 CURRENT APPLICATION NUMBER: US/09/7796,487
 CURRENT FILING DATE: 2001-03-02
 PRIOR APPLICATION NUMBER: US 60/186,532
 PRIOR FILING DATE: 2000-03-05
 PRIOR APPLICATION NUMBER: US 09/530,868
 PRIOR FILING DATE: 2000-05-05

```
;; TITLE OF INVENTION: REGULATORS OF INTRACELLULAR PHOSPHORYLATION
;; FILE REFERENCE: PF-0683 P
;; CURRENT APPLICATION NUMBER: US/60/125,593
;; CURRENT FILING DATE: 1999-03-18
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PERL Program
;; SEQ ID NO 5
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: 2415617
US-60-125-593-5

Query Match
Best Local Similarity 100.0%; Score 2016; DB 27; Length 384;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDPAGGPRGVLPKPCRVLYLNLNPRGKGKALQLFPSHVOPPLAEAEISFTLMLTERRNHA 60
DB 1 MDPAGGPRGVLPKPCRVLYLNLNPRGKGKALQLFPSHVOPPLAEAEISFTLMLTERRNHA 60
OY 61 RELVSEELGRMDALVWMSGDGLMHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVWMSGDGLMHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASL 120
OY 121 NHYAGEQVNTNEDLLTNCTLLCRRLSPMNLISHTASGLRFLSVLSIAMGFIDVDLE 180
DB 121 NHYAGEQVNTNEDLLTNCTLLCRRLSPMNLISHTASGLRFLSVLSIAMGFIDVDLE 180
OY 181 SEKYRRLGEMRFTLTGTFRLAALRTYRGLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
DB 181 SEKYRRLGEMRFTLTGTFRLAALRTYRGLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
OY 241 LEEPPSHMTVVPDDEDFVLYLALLSHLSGSEMFAPMGCAGVNHLEFYRAGVSRAML 300
DB 241 LEEPPSHMTVVPDDEDFVLYLALLSHLSGSEMFAPMGCAGVNHLEFYRAGVSRAML 300
OY 301 RLFLAMEKGRHMEYECPLYVYVVAFLRLEPKDGKGFVAVDELAVSEAVOQOVHPNIFW 360
DB 301 RLFLAMEKGRHMEYECPLYVYVVAFLRLEPKDGKGFVAVDELAVSEAVOQOVHPNIFW 360
OY 361 MVSQGVPEPPPSMKPQOMPPPEEPL 384
DB 361 MVSQGVPEPPPSMKPQOMPPPEEPL 384

RESULT 7
US-09-959-897-2
;; Sequence 2, Application US/09959897
;; GENERAL INFORMATION:
;; APPLICANT: PITSON, Stuart M
;; APPLICANT: Brian, MATTEMBERG W
;; APPLICANT: Pu, XIA
;; APPLICANT: Richard, D'ANDREA J
;; APPLICANT: Jennifer, BAMBLE R
;; APPLICANT: Mathew, VADAS A
;; TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME
;; FILE REFERENCE: PITSON-1
;; CURRENT FILING DATE: 2001-11-13
;; PRIOR APPLICATION NUMBER: US/09/959, 897
;; PRIOR FILING DATE: 2000-05-12
;; PRIOR APPLICATION NUMBER: PCT/AU00/00457
;; PRIOR FILING DATE: 1999-05-13
;; PRIOR APPLICATION NUMBER: AU PO 0339
;; PRIOR FILING DATE: 1999-07-08
;; PRIOR APPLICATION NUMBER: AU PO 1504
;; NUMBER OF SEQ ID NOS: 56
;; SOFTWARE: PatentIn, version 3.1
;; SEQ ID NO 2
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo sapiens
```

```
US-09-959-897-2

Query Match
Best Local Similarity 99.9%; Score 2013; DB 23; Length 384;
Matches 383; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDPAGGPRGVLPKPCRVLYLNLNPRGKGKALQLFPSHVOPPLAEAEISFTLMLTERRNHA 60
DB 1 MDPAGGPRGVLPKPCRVLYLNLNPRGKGKALQLFPSHVOPPLAEAEISFTLMLTERRNHA 60
OY 61 RELVSEELGRMDALVWMSGDGLMHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVWMSGDGLMHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASL 120
OY 121 NHYAGEQVNTNEDLLTNCTLLCRRLSPMNLISHTASGLRFLSVLSIAMGFIDVDLE 180
DB 121 NHYAGEQVNTNEDLLTNCTLLCRRLSPMNLISHTASGLRFLSVLSIAMGFIDVDLE 180
OY 181 SEKYRRLGEMRFTLTGTFRLAALRTYRGLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
DB 181 SEKYRRLGEMRFTLTGTFRLAALRTYRGLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
OY 241 LEEPPSHMTVVPDDEDFVLYLALLSHLSGSEMFAPMGCAGVNHLEFYRAGVSRAML 300
DB 241 LEEPPSHMTVVPDDEDFVLYLALLSHLSGSEMFAPMGCAGVNHLEFYRAGVSRAML 300
OY 301 RLFLAMEKGRHMEYECPLYVYVVAFLRLEPKDGKGFVAVDELAVSEAVOQOVHPNIFW 360
DB 301 RLFLAMEKGRHMEYECPLYVYVVAFLRLEPKDGKGFVAVDELAVSEAVOQOVHPNIFW 360
OY 361 MVSQGVPEPPPSMKPQOMPPPEEPL 384
DB 361 MVSQGVPEPPPSMKPQOMPPPEEPL 384

RESULT 8
US-60-364-517-2
;; Sequence 2, Application US/60364517
;; GENERAL INFORMATION:
;; APPLICANT: Tsai, Pong-Ying
;; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;; FILE REFERENCE: MPIO2-032P1 (M)
;; CURRENT APPLICATION NUMBER: US/60/364,517
;; CURRENT FILING DATE: 2002-03-15
;; NUMBER OF SEQ ID NOS: 2
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-60-364-517-2

Query Match
Best Local Similarity 99.9%; Score 2013; DB 27; Length 384;
Matches 383; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDPAGGPRGVLPKPCRVLYLNLNPRGKGKALQLFPSHVOPPLAEAEISFTLMLTERRNHA 60
DB 1 MDPAGGPRGVLPKPCRVLYLNLNPRGKGKALQLFPSHVOPPLAEAEISFTLMLTERRNHA 60
OY 61 RELVSEELGRMDALVWMSGDGLMHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVWMSGDGLMHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASL 120
OY 121 NHYAGEQVNTNEDLLTNCTLLCRRLSPMNLISHTASGLRFLSVLSIAMGFIDVDLE 180
DB 121 NHYAGEQVNTNEDLLTNCTLLCRRLSPMNLISHTASGLRFLSVLSIAMGFIDVDLE 180
OY 181 SEKYRRLGEMRFTLTGTFRLAALRTYRGLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
DB 181 SEKYRRLGEMRFTLTGTFRLAALRTYRGLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
```

OTHER INFORMATION: Incyte ID No: 2415617CD1
US-09-937-060A-5

Query Match 100.0%; Score 2016; DB 23; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.2e-203;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGPRGVLPKRCRYLYLNPBGKGAOLFRSHVOPPLAESEFTLMTERRNHA 60
DB 1 MDPAGGPRGVLPKRCRYLYLNPBGKGAOLFRSHVOPPLAESEFTLMTERRNHA 60
QY 61 RELVSEELGRMDALVWVSGDGLMHEVYVNGLMERPDWETAIOKPLCSIPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVWVSGDGLMHEVYVNGLMERPDWETAIOKPLCSIPAGSGNALAASL 120
QY 121 NHYAGYEQVNTNEDLLTNCITLLCRLLSPNNLLSHITASGLRFSVLSLWAGFIADVDLE 180
DB 121 NHYAGYEQVNTNEDLLTNCITLLCRLLSPNNLLSHITASGLRFSVLSLWAGFIADVDLE 180
QY 181 SEKVRRLGEMRFTLTGTFRLAALRTYGRLAYLPVGRVSKTPASPVVOOGCPVDALHVP 240
DB 181 SEKVRRLGEMRFTLTGTFRLAALRTYGRLAYLPVGRVSKTPASPVVOOGCPVDALHVP 240
QY 241 LEEPVPSHMTVVPDEDFVLYLALSHLSEMFAPMGRCAAGVMHLFYRAGVSRAML 300
DB 241 LEEPVPSHMTVVPDEDFVLYLALSHLSEMFAPMGRCAAGVMHLFYRAGVSRAML 300
QY 301 RFLAMEGRHMEYECPLYVYVVAFRLEPKDKGVFAVDGELMVSEAVOGVHPNFTW 360
DB 301 RFLAMEGRHMEYECPLYVYVVAFRLEPKDKGVFAVDGELMVSEAVOGVHPNFTW 360
QY 361 MVSQGVPEPPPSMKPOQMPPEEPL 384
DB 361 MVSQGVPEPPPSMKPOQMPPEEPL 384

RESULT 4

US-09-970-516-2
Sequence 2, Application US/09970516
GENERAL INFORMATION:
APPLICANT: Novartis AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-31617
CURRENT APPLICATION NUMBER: US/09/970,516
CURRENT FILING DATE: 2001-10-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
US-09-970-516-2

Query Match 100.0%; Score 2016; DB 23; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.2e-203;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGPRGVLPKRCRYLYLNPBGKGAOLFRSHVOPPLAESEFTLMTERRNHA 60
DB 1 MDPAGGPRGVLPKRCRYLYLNPBGKGAOLFRSHVOPPLAESEFTLMTERRNHA 60
QY 61 RELVSEELGRMDALVWVSGDGLMHEVYVNGLMERPDWETAIOKPLCSIPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVWVSGDGLMHEVYVNGLMERPDWETAIOKPLCSIPAGSGNALAASL 120
QY 121 NHYAGYEQVNTNEDLLTNCITLLCRLLSPNNLLSHITASGLRFSVLSLWAGFIADVDLE 180
DB 121 NHYAGYEQVNTNEDLLTNCITLLCRLLSPNNLLSHITASGLRFSVLSLWAGFIADVDLE 180
QY 181 SEKVRRLGEMRFTLTGTFRLAALRTYGRLAYLPVGRVSKTPASPVVOOGCPVDALHVP 240
DB 181 SEKVRRLGEMRFTLTGTFRLAALRTYGRLAYLPVGRVSKTPASPVVOOGCPVDALHVP 240

QY 241 LEEPVPSHMTVVPDEDFVLYLALSHLSEMFAPMGRCAAGVMHLFYRAGVSRAML 300
DB 241 LEEPVPSHMTVVPDEDFVLYLALSHLSEMFAPMGRCAAGVMHLFYRAGVSRAML 300
QY 301 RFLAMEGRHMEYECPLYVYVVAFRLEPKDKGVFAVDGELMVSEAVOGVHPNFTW 360
DB 301 RFLAMEGRHMEYECPLYVYVVAFRLEPKDKGVFAVDGELMVSEAVOGVHPNFTW 360
QY 361 MVSQGVPEPPPSMKPOQMPPEEPL 384
DB 361 MVSQGVPEPPPSMKPOQMPPEEPL 384

RESULT 5

US-10-111-729-3
Sequence 3, Application US/10111729
GENERAL INFORMATION:
APPLICANT: Warner-lambert
TITLE OF INVENTION: Human sphingosine kinase gene
FILE REFERENCE: A0001986CT
CURRENT APPLICATION NUMBER: US/10/111,729
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
US-10-111-729-3

Query Match 100.0%; Score 2016; DB 25; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.2e-203;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGPRGVLPKRCRYLYLNPBGKGAOLFRSHVOPPLAESEFTLMTERRNHA 60
DB 1 MDPAGGPRGVLPKRCRYLYLNPBGKGAOLFRSHVOPPLAESEFTLMTERRNHA 60
QY 61 RELVSEELGRMDALVWVSGDGLMHEVYVNGLMERPDWETAIOKPLCSIPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVWVSGDGLMHEVYVNGLMERPDWETAIOKPLCSIPAGSGNALAASL 120
QY 121 NHYAGYEQVNTNEDLLTNCITLLCRLLSPNNLLSHITASGLRFSVLSLWAGFIADVDLE 180
DB 121 NHYAGYEQVNTNEDLLTNCITLLCRLLSPNNLLSHITASGLRFSVLSLWAGFIADVDLE 180
QY 181 SEKVRRLGEMRFTLTGTFRLAALRTYGRLAYLPVGRVSKTPASPVVOOGCPVDALHVP 240
DB 181 SEKVRRLGEMRFTLTGTFRLAALRTYGRLAYLPVGRVSKTPASPVVOOGCPVDALHVP 240
QY 241 LEEPVPSHMTVVPDEDFVLYLALSHLSEMFAPMGRCAAGVMHLFYRAGVSRAML 300
DB 241 LEEPVPSHMTVVPDEDFVLYLALSHLSEMFAPMGRCAAGVMHLFYRAGVSRAML 300
QY 301 RFLAMEGRHMEYECPLYVYVVAFRLEPKDKGVFAVDGELMVSEAVOGVHPNFTW 360
DB 301 RFLAMEGRHMEYECPLYVYVVAFRLEPKDKGVFAVDGELMVSEAVOGVHPNFTW 360
QY 361 MVSQGVPEPPPSMKPOQMPPEEPL 384
DB 361 MVSQGVPEPPPSMKPOQMPPEEPL 384

RESULT 6

US-60-125-593-5
Sequence 5, Application US/60125593
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Tang, Y. Tom
APPLICANT: Yue, Henry
APPLICANT: Baughn, Mariah R.
APPLICANT: Azimzai, Yalda

```

; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 13997
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-629-469A-13997

```

```

Query Match          100.0%; Score 2016; DB 20; Length 384;
Best Local Similarity 100.0%; Pred. No. 1,2e-203;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGPRGVLPKPCRVLYLNLPRGGKALQLFRSHVQPLLAEAISFTMLTERRNHA 60
DB 1 MDPAGGPRGVLPKPCRVLYLNLPRGGKALQLFRSHVQPLLAEAISFTMLTERRNHA 60
QY 61 RELVSEELGRMDALVVMGSDGLMHEVYNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVVMGSDGLMHEVYNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
QY 121 NHYAGEQVNTNEDLTNCTLLCRLLSPMNLISLFTASGLRFSVLSLAWGFADVDLE 180
DB 121 NHYAGEQVNTNEDLTNCTLLCRLLSPMNLISLFTASGLRFSVLSLAWGFADVDLE 180
QY 181 SEKRYRRLGEMRRTLTGTFRLAALRTYRGLAYLPVGRVSGSKTPASPVVVOGQPVDAHLVP 240
DB 181 SEKRYRRLGEMRRTLTGTFRLAALRTYRGLAYLPVGRVSGSKTPASPVVVOGQPVDAHLVP 240
QY 241 LEEPVSHTVVPDDEDFVLYLALLSHLSGSEMPFAAPMGCAAGVNHLPFVRAGVSRAML 300
DB 241 LEEPVSHTVVPDDEDFVLYLALLSHLSGSEMPFAAPMGCAAGVNHLPFVRAGVSRAML 300
QY 301 RLFLAMEKGRHMEYECPLYVYVVAFAFLRLEPKDGKGFVAVDGLMVSSEAVOGQVHPNRYM 360
DB 301 RLFLAMEKGRHMEYECPLYVYVVAFAFLRLEPKDGKGFVAVDGLMVSSEAVOGQVHPNRYM 360
QY 361 MVSGCEVPEPPSMKPOOMPPEEPL 384
DB 361 MVSGCEVPEPPSMKPOOMPPEEPL 384

```

```

RESULT 2
US-09-629-469A-15394
; Sequence 15394, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTO, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899

```

```

; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 15394
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-629-469A-15394

```

```

Query Match          100.0%; Score 2016; DB 20; Length 384;
Best Local Similarity 100.0%; Pred. No. 1,2e-203;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGPRGVLPKPCRVLYLNLPRGGKALQLFRSHVQPLLAEAISFTMLTERRNHA 60
DB 1 MDPAGGPRGVLPKPCRVLYLNLPRGGKALQLFRSHVQPLLAEAISFTMLTERRNHA 60
QY 61 RELVSEELGRMDALVVMGSDGLMHEVYNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVVMGSDGLMHEVYNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
QY 121 NHYAGEQVNTNEDLTNCTLLCRLLSPMNLISLFTASGLRFSVLSLAWGFADVDLE 180
DB 121 NHYAGEQVNTNEDLTNCTLLCRLLSPMNLISLFTASGLRFSVLSLAWGFADVDLE 180
QY 181 SEKRYRRLGEMRRTLTGTFRLAALRTYRGLAYLPVGRVSGSKTPASPVVVOGQPVDAHLVP 240
DB 181 SEKRYRRLGEMRRTLTGTFRLAALRTYRGLAYLPVGRVSGSKTPASPVVVOGQPVDAHLVP 240
QY 241 LEEPVSHTVVPDDEDFVLYLALLSHLSGSEMPFAAPMGCAAGVNHLPFVRAGVSRAML 300
DB 241 LEEPVSHTVVPDDEDFVLYLALLSHLSGSEMPFAAPMGCAAGVNHLPFVRAGVSRAML 300
QY 301 RLFLAMEKGRHMEYECPLYVYVVAFAFLRLEPKDGKGFVAVDGLMVSSEAVOGQVHPNRYM 360
DB 301 RLFLAMEKGRHMEYECPLYVYVVAFAFLRLEPKDGKGFVAVDGLMVSSEAVOGQVHPNRYM 360
QY 361 MVSGCEVPEPPSMKPOOMPPEEPL 384
DB 361 MVSGCEVPEPPSMKPOOMPPEEPL 384

```

```

RESULT 3
US-09-937-060A-5
; Sequence 5, Application US/09937060A
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: BANDMAN, OLGA
; APPLICANT: TANG, Y. TOM
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Dying Aina M.
; APPLICANT: AU-YOUNG, Janice
; TITLE OF INVENTION: REGULATORS OF INTRACELLULAR PHOSPHORYLATION
; FILE REFERENCE: PF-0683 PCT
; CURRENT APPLICATION NUMBER: US/09/937,060A
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: 60/125,593; 60/135,049; 60/143,188
; PRIOR FILING DATE: 1999-03-18; 1999-05-20; 1999-07-09
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:46:23 : Search time 149 Seconds
(without alignments)
1061.594 Million cell updates/sec

Title: US-09-937-060A-5
Perfect score: 2016
Sequence: 1 MDPAGPGRVLPKRCVILV.....CPEPPSMKPDMPPEEPL 384

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 4569144 segs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCYUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US103_COMB.pep:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2016	100.0	384	20	US-09-629-469A-13997
2	2016	100.0	384	20	US-09-629-469A-15394
3	2016	100.0	384	23	US-09-937-060A-5
4	2016	100.0	384	23	US-09-970-516-2
5	2016	100.0	384	25	US-10-111-729-3
6	2016	100.0	384	27	US-60-125-593-5

7	2013	99.9	384	23	US-09-959-897-2	Sequence 2, App11
8	2013	99.9	384	27	US-60-364-517-2	Sequence 2, App11
9	2011	99.8	384	18	US-09-488-725A-3325	Sequence 3325, Ap
10	2008	99.6	384	1	PCT-US02-24623-10	Sequence 10, App1
11	2008	99.6	384	21	US-09-796-487-3	Sequence 3, App11
12	1962	97.3	384	21	US-09-784-810A-2	Sequence 2, App11
13	1923	95.4	368	24	US-10-053-510-21	Sequence 21, App1
14	1752	86.9	333	1	PCT-US01-16450-2585	Sequence 2585, Ap
15	1732	86.9	333	1	PCT-US01-16450A-2585	Sequence 2585, Ap
16	1639.5	81.3	382	21	US-09-784-810A-4	Sequence 4, App11
17	1632.5	81.0	382	23	US-09-970-516-6	Sequence 6, App11
18	1630.5	80.9	388	1	PCT-US01-09664A-15	Sequence 15, App11
19	1630.5	80.9	388	21	US-09-796-487-2	Sequence 2, App11
20	1630.5	80.9	388	22	US-09-817-676A-15	Sequence 15, App11
21	1627.5	80.7	381	21	US-09-796-487-1	Sequence 4, App11
22	1627.5	80.7	381	21	US-09-796-487-4	Sequence 4, App11
23	1611	79.9	305	18	US-09-488-725A-6897	Sequence 6897, Ap
24	1609.5	79.8	373	21	US-09-796-487-5	Sequence 5, App11
25	1541	76.4	293	1	PCT-US01-05614-328	Sequence 328, App
26	1541	76.4	293	16	US-09-205-258-328	Sequence 313, App
27	1541	76.4	293	16	US-09-205-258-328	Sequence 328, App
28	1541	76.4	293	23	US-09-933-767-328	Sequence 328, App
29	1541	76.4	293	24	US-10-004-860-328	Sequence 328, App
30	1541	76.4	293	24	US-10-023-282-328	Sequence 328, App
31	1477.5	73.3	381	27	US-60-243-468-1133	Sequence 1133, Ap
32	1077.5	53.4	260	20	US-09-641-377-255	Sequence 255, App
33	940	46.6	548	27	US-60-243-468-719	Sequence 719, App
34	938.5	46.6	204	21	US-09-796-487-9	Sequence 9, App11
35	893	44.3	617	1	PCT-US01-09664A-12	Sequence 12, App1
36	893	44.3	617	22	US-09-817-676A-12	Sequence 12, App1
37	886.5	44.0	618	1	PCT-US01-09664A-14	Sequence 14, App1
38	886.5	44.0	618	1	PCT-US02-24623-11	Sequence 11, App1
39	886.5	44.0	618	22	US-09-817-676A-14	Sequence 14, App1
40	886.5	44.0	618	22	US-09-970-516-4	Sequence 4, App11
41	886.5	44.0	618	27	US-60-388-853-4	Sequence 254, App
42	886.5	44.0	654	20	US-09-641-377-254	Sequence 254, App
43	873.5	43.3	806	1	PCT-US01-08631-51503	Sequence 51503, A
44	833.5	41.3	638	23	US-09-969-896-3	Sequence 3, App11
45	694	34.4	146	27	US-60-178-307-3020	Sequence 3020, Ap

ALIGNMENTS

RESULT 1
US-09-629-469A-13997
Sequence 13997 Application US/09629469A
GENERAL INFORMATION:
APPLICANT: OYA, TOSHIO
APPLICANT: ISOGAI, TAKAO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: HAYASHI, KOJI
APPLICANT: SAITO, KAORU
APPLICANT: YAMAMOTO, JUNICHI
APPLICANT: ISHII, SHIZUKO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: WAKAMATSU, AI
APPLICANT: NAGAI, KEIICHI
APPLICANT: OTSUKI, TETSUJI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
FILE REFERENCE: 084335/0123
CURRENT APPLICATION NUMBER: US/09/629, 469A
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: JP 1999-248036
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: JP 1999-300253
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09

; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-784-810a-15

Query Match 18.68; Score 374.5; DB 10; Length 415;
Best Local Similarity 28.68; Pred. No. 2,9e-28;
Matches 110; Conservative 80; Mismatches 146; Indels 49; Gaps 12;

```

QY 12 PRP---CR--VLVLINPRGKGKALOLFRRSHVOPFLAEAEISFTLMTERRNHARE-LVR 65
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 20 PPPOEOCRGNILVFINNSGTGKSLFTFANTVCGPKLDKSLIRYEVVYTTGPNHARNVLTMT 79
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 66 SEELGRMDALVMSGDIIMEVNGIMERP-DWETAIOKPLCSLPAGSGNALAASLNHYA 124
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 80 KADLGKFNGLVILSGDGLVPEALNGILCREDAFRIPTLPIGIVPSGSGNGLCSV--LS 137
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 125 GYEQVTNEDLITNCTLLCLRRLSPMNLSLHT--ASGLRFSVLSLAWGFADVDLESE 182
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 138 KYGTMMNEKSVMERALEIATSPTKAESVALYSVKTNDQSYASFSLISGWMADIDIDSE 197
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 183 KYRR-LGEMRPTLTGFLRLAALFTYRGRLAYLP-----VGRYSKTPASPV--- 227
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 198 KWRKSLGHHRTVMGFIKSCNLRSTYKGRILYRYPKGFHRSSNVFSYEXTTQQRIDDS 257
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 228 -VVOQGPDVAHLVPLEPVPSSH-----TVVPDEDEFLVLLLSHL 268
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 258 KVKTNQSVSDEEETMETKFCQNTLIPDSDETLAVGSSDLEETVVIDNPNVNIYAVTLSHI 317
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 269 GSEMFAPMGRCACGVNHLFTV--RAGVSRMLRLFLAMKGRHMEYECPLYVYPVYA 326
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 318 AADGPFAPASAKLEDNRILSLIIMKIDIGTRVNIKAYLLAIEHETHL--DLPFVAKHEVSS 375
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 327 FRLEPKDQKGVFAVDELMVSEAVQ 351
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 376 MKLEVISRSGSHVLDGEVVDKTIIE 400
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: May 9, 2003, 16:54:15
Job time : 39 secs

Db 616 Q 616

US-09-817-676A-14
Sequence 14 Application US/09817676A
Patent No. US20020042101A1

GENERAL INFORMATION:
APPLICANT: Spigel, Sarah
APPLICANT: Kohama, Takatumi
TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
FILE REFERENCE: 00170/HG
CURRENT APPLICATION NUMBER: US/09/817,676A
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/194,318
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 618
TYPE: PRT
ORGANISM: Homo sapiens
US-09-817-676A-14

Query Match

Best Local Similarity 44.0%; Score 886.5; DB 10; Length 618;
Matches 195; Conservative 61; Mismatches 107; Indels 127; Gaps 6;

QY 10 VLPKRCVLYLNLNPGCKKALQLFPSHVOPPLAELISFTLMITERRNHARELVSEEL 69
DB 140 LLPKRPRLLLVNPFGGRLAWQCKNHVLPMSISAGLSFTLIQTERQNHARELVQGLSL 199
QY 70 GWRDALVYMSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAGSGNALAASLNHYAGYEV 129
DB 200 SEMDGIIVYSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAGSGNALAASLNHYAGYEV 259
QY 130 TNEDELITNCTLLCRLLSPMNLISLTASGLRFSVSLAMGFIADVDESERYRLGE 189
DB 260 LQIDLILNCSLLICRGCGHPLDLISVTLASGSRCSFSLVAMGFSDVDIQRFRALGS 319
QY 190 MREFTLGTFLRLAALRYRGRALYPVG-RVGSKTPA----- 224
DB 320 ABEFTLGTGLATLHTYRGRSLYPATVBPASPTPAHSLPRAKSELTLTPDPAPMANSP 379
QY 225 ----- 224
DB 380 LHRVSDDLPLPQALASPGSEPLPLISLNGGRELADMGAGADAPLSPDLLSSPP 439
QY 225 -----SPVYVQO-----GPVDALVPLEEPVSHWTVP 253
DB 440 GSPKAAHSPVSEGAIVIPSSGLPLPTPDARVAGSTCGPRDHLPLPLGTPPLPDDM-VTL 498
QY 254 DDEFLVYALHLSHLSSEMFARPMGRCAAGVNHLYFVAGVSRAMLLRLFLAMEKGRHME 313
DB 499 EEDFLVLMALISPSHLGADVAARHAFDDGLVHLICWVSGISRAALLRLFLAMERSHFS 558
QY 314 YECPRLYVVPVAFRLPKDKGVAFVADGELMVSFAVGVNHYFMVSGCVPEPPSMK 373
DB 559 LCCPOLGYAARARLEPLRPGVLTVDGEVEYGRLOAOMHNGIGTLTG-----PPGC- 613
QY 374 PQQMPPEEP 383
DB 614 -----PGREP 618

RESULT 14
US-09-970-516-4
Sequence 4 Application US/09970516
Patent No. US20020099029A1
GENERAL INFORMATION:
APPLICANT: NO. US20020099029A1arcls AG

TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: A-31617
CURRENT APPLICATION NUMBER: US/09/970,516
CURRENT FILING DATE: 2001-10-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 618
TYPE: PRT
ORGANISM: Homo sapiens
US-09-970-516-4

Query Match 44.0%; Score 886.5; DB 10; Length 618;
Best Local Similarity 39.8%; Pred. No. 4.3e-78;
Matches 195; Conservative 61; Mismatches 107; Indels 127; Gaps 6;

QY 10 VLPKRCVLYLNLNPGCKKALQLFPSHVOPPLAELISFTLMITERRNHARELVSEEL 69
DB 140 LLPKRPRLLLVNPFGGRLAWQCKNHVLPMSISAGLSFTLIQTERQNHARELVQGLSL 199
QY 70 GWRDALVYMSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAGSGNALAASLNHYAGYEV 129
DB 200 SEMDGIIVYSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAGSGNALAASLNHYAGYEV 259
QY 130 TNEDELITNCTLLCRLLSPMNLISLTASGLRFSVSLAMGFIADVDESERYRLGE 189
DB 260 LQIDLILNCSLLICRGCGHPLDLISVTLASGSRCSFSLVAMGFSDVDIQRFRALGS 319
QY 190 MREFTLGTFLRLAALRYRGRALYPVG-RVGSKTPA----- 224
DB 320 ABEFTLGTGLATLHTYRGRSLYPATVBPASPTPAHSLPRAKSELTLTPDPAPMANSP 379
QY 225 ----- 224
DB 380 LHRVSDDLPLPQALASPGSEPLPLISLNGGRELADMGAGADAPLSPDLLSSPP 439
QY 225 -----SPVYVQO-----GPVDALVPLEEPVSHWTVP 253
DB 440 GSPKAAHSPVSEGAIVIPSSGLPLPTPDARVAGSTCGPRDHLPLPLGTPPLPDDM-VTL 498
QY 254 DDEFLVYALHLSHLSSEMFARPMGRCAAGVNHLYFVAGVSRAMLLRLFLAMEKGRHME 313
DB 499 EEDFLVLMALISPSHLGADVAARHAFDDGLVHLICWVSGISRAALLRLFLAMERSHFS 558
QY 314 YECPRLYVVPVAFRLPKDKGVAFVADGELMVSFAVGVNHYFMVSGCVPEPPSMK 373
DB 559 LCCPOLGYAARARLEPLRPGVLTVDGEVEYGRLOAOMHNGIGTLTG-----PPGC- 613
QY 374 PQQMPPEEP 383
DB 614 -----PGREP 618

RESULT 15
US-09-784-810A-15
Sequence 15 Application US/09784810A
Patent No. US20020082203A1
GENERAL INFORMATION:
APPLICANT: RASTELLI, LUCA
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
FILE REFERENCE: 10716-08
CURRENT APPLICATION NUMBER: US/09/784,810A
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/182,360
PRIOR FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 415


```

QY 10 VLPRRCRVLLNPRGKGKALQIFRSHVQPLLAELISFTLMLTERNNHARELVSEEL 69
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LPRRCRVLLNPRGKGKALQIFRSHVQPLLAELISFTLMLTERNNHARELVSEEL 60
QY 70 GRMDALVMSGDGLMHEVNGMLERPDWETAIOKPLCSLPAGSGNALAASLNHYAGYEQY 129
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GRMDALVMSGDGLMHEVNGMLERPDWETAIOKPLCSLPAGSGNALAASLNHYAGYEQY 120
QY 130 TNEEDLLTNCITLLCRRLSPMNLISLHTASGLRIFSVLSLAWGFIADVDLESEKRRRLGE 189
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 TNEEDLLTNCITLLCRRLSPMNLISLHTASGLRIFSVLSLAWGFIADVDLESEKRRRLGE 180
QY 190 MRFTLGTFRLALARTYRGRLAYLPVGRVSKTPASPVVVOGPDVAHLVPLEEVPSSH 249
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 IRFTYGTFFRLASRLTYGQGLALCPVGTIVASKRPAS-TLYOKGPDVDFHLVPLEEVPSSH 239
QY 230 TVPDEDEVLVYALHSHLSEMPAAMPGRCAAGVNHLYFYVAGYSKRAMLLFLAMEKG 309
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 TVPDEDEVLVYALHSHLSEMPAAMPGRCAAGVNHLYFYVAGYSKRAMLLFLAMEKG 299
QY 310 RHMEYECPLYVVPVAVRLEPKDGKGFVAVDGLMWSAEOGVHPNYFMVSGCPEPP 369
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 RHMEYECPLYVVPVAVRLEPKDGKGFVAVDGLMWSAEOGVHPNYFMVSGCPEPP 359
QY 370 PSMKPOAMPPEEP 383
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 SGRDSRGRPPPEEP 373

```

RESULT 11

US-09-796-487-9

Sequence 9, Application US/09796487

Patent No. US20020042358A1

GENERAL INFORMATION:

APPLICANT: Spiegel, Sarah

TITLE OF INVENTION: Splicing Kinase, Cloning, Expression and Methods of Use

FILE REFERENCE: 07320001a (2033957-0001)

CURRENT APPLICATION NUMBER: US/09/796,487

CURRENT FILING DATE: 2001-03-02

PRIOR APPLICATION NUMBER: US 60/186,532

PRIOR FILING DATE: 2000-03-03

PRIOR APPLICATION NUMBER: US 09/530,868

PRIOR FILING DATE: 2000-05-05

NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn version 3.1

SEQ ID NO 9

LENGTH: 204

TYPE: PRT

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: Putative kinase sequence obtained by assembling sequences from

OTHER INFORMATION: several human ESTs (accession numbers D31133, AA232791, W63556, A

NAME/KEY: MISC_FEATURE

LOCATION: (1)..(204)

OTHER INFORMATION: Corresponding to peptide sequence Putative Human in Figure 2.

US-09-796-487-9

```

Query Match 46.6%; Score 938.5; DB 10; Length 204;
Best local Similarity 58.4%; Pred. No. 7.4e-84;
Matches 199; Conservative 3; Mismatches 2; Indels 137; Gaps 2;
QY 21 LNRGKGKALQIFRSHVQPLLAELISFTLMLTERNNHARELVSEELGRMDALVMSG 80
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LNRGKGKALQIFRSHVQPLLAELISFTLMLTERNNHARELVSEELGRMDALVMSG 60
QY 81 DGLMHEVNGMLERPDWETAIOKPLCSLPAGSGNALAASLNHYAGYEQY TNEEDLLNCT 140
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 DGLMHEVNGMLERPDWETAIOKPLCSLPAGSGNALAASLNHYAGYEQY TNEEDLLNCT 107
QY 141 LCRRLSPMNLISLHTASGLRIFSVLSLAWGFIADVDLESEKRRRLGEFTLRL 200
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 108 LCRRLSPMNLISLHTASGLRIFSVLSLAWGFIADVDLESEKRRRLGEFTLRL 145

```

```

QY 201 AALRTYRGRLAYLPVGRVSKTPASPVVVOGPDVAHLVPLEEVPSSHMTVPDEDFVLV 260
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 146 AALRTYRGRLAYLPVGRVSKTPASPVVVOGPDVAHLVPLEEVPSSHMTVPDEDFVLV 165
QY 261 LALHSHLSEMPAAMPGRCAAGVNHLYFYVAGYSKRAMLLFLAMEKGRHMEYECPLY 320
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 167 LALHSHLSEMPAAMPGRCAAGVNHLYFYVAGYSKRAMLLFLAMEKGRHMEYECPLY 166
QY 321 YVPVAVRLEPKDGKGFVAVDGLMWSAEOGVHPNYFMVSGCPEPP 361
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 167 YVPVAVRLEPKDGKGFVAVDGLMWSAEOGVHPNYFMVSGCPEPP 204

```

RESULT 12

US-09-817-676A-12

Sequence 12, Application US/09817676A

Patent No. US20020042101A1

GENERAL INFORMATION:

APPLICANT: Spiegel, Sarah

TITLE OF INVENTION: Mammalian Splicing Kinase Type 2 Isoforms, Cloning,

TITLE OF INVENTION: Expression and Methods of Use Thereof

FILE REFERENCE: 00170/HG

CURRENT APPLICATION NUMBER: US/09/817,676A

CURRENT FILING DATE: 2001-03-26

PRIOR APPLICATION NUMBER: US 60/194,318

PRIOR FILING DATE: 2000-04-03

NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 12

LENGTH: 617

TYPE: PRT

ORGANISM: Mus musculus

US-09-817-676A-12

```

Query Match 44.3%; Score 893; DB 10; Length 617;
Best local Similarity 39.9%; Pred. No. 9.9e-79;
Matches 192; Conservative 61; Mismatches 108; Indels 120; Gaps 5;
QY 10 VLPRRCRVLLNPRGKGKALQIFRSHVQPLLAELISFTLMLTERNNHARELVSEEL 69
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 141 LPRRCRVLLNPRGKGKALQIFRSHVQPLLAELISFTLMLTERNNHARELVSEEL 200
QY 70 GRMDALVMSGDGLMHEVNGMLERPDWETAIOKPLCSLPAGSGNALAASLNHYAGYEQY 129
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 201 SEMEGIVTAVSGDGLMHEVNGMLERPDWETAIOKPLCSLPAGSGNALAASLNHYAGYEQY 260
QY 130 TNEEDLLTNCITLLCRRLSPMNLISLHTASGLRIFSVLSLAWGFIADVDLESEKRRRLGE 189
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 261 VGYDILLNCSLILCRGSHPLDLLSVTLASGSRCSFSLVAMGFLSDVDIHSEFRALGS 320
QY 190 MRFTLGTFRLALARTYRGRLAYLPVGRVSKTPASPVVVOGPDVAHLVPLEEVPSSH 249
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 321 ARFTLGAVGLASLHRTYRGRLAYLPVGRVSKTPASPVVVOGPDVAHLVPLEEVPSSH 380
QY 216 ARFTLGAVGLASLHRTYRGRLAYLPVGRVSKTPASPVVVOGPDVAHLVPLEEVPSSH 220
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 381 LHRVSQPLPLPQALVSPSPPEPLDLSLNGGPELTGDMGAGCAPLSPDPLPSSP 440
QY 221 LHRVSQPLPLPQALVSPSPPEPLDLSLNGGPELTGDMGAGCAPLSPDPLPSSP 254
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 441 NALKTQSLPIAGPMPMPASGFLPETHSAPASTWGPDHLLPLGSPLPDQWWTIEG 500
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 255 EDEVLVYALHSHLSEMPAAMPGRCAAGVNHLYFYVAGYSKRAMLLFLAMEKGRHMEY 314
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 501 E-FLYLMGILTSILCADLMAPIARPDGVLHLCWVNSGSLRALLRIFLAMEGNHFS 529
QY 315 ECPYLVVVPVAVRLEPKDGKGFVAVDGLMWSAEOGVHPNYFMVSGCPEPPSWKP 374
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 560 GCPHLYAARAARFLEPLTRGLTGVDELVEYGPIDQAVHPGLATILTG-PAGOKP 615
QY 375 GCPHLYAARAARFLEPLTRGLTGVDELVEYGPIDQAVHPGLATILTG-PAGOKP 375

```

Best Local Similarity 81.4%; Pred. No. 7,4e-151;
Matches 307; Conservative 33; Mismatches 36; Indels 1; Gaps 1;

```
QY 7 PGCVLPKPCRYLVLLNPGKGRKALQLFRSHVOPLLAEAEISFTLMTERRNNHARELVRS 66
Db 6 PRLGLRPRCRYLVLLNPGKGRKALQLFRSHVOPFLPEAEITFKILTERKNHARELVCA 65
QY 67 EELGRMDALVYVSGDGLMHEVYVNGLMERPMETAIQKPLCSIPAGSGNALAASVNHAGY 126
Db 66 EELGHMDALVYVSGDGLMHEVYVNGLMERPMETAIQKPLCSIPAGSGNALAASVNHAGY 125
QY 127 EGYTNEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSLACGFIADVDLESEKYYR 186
Db 126 EGYTNEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSLACGFIADVDLESEKYYR 185
QY 187 LCGMRFTLGTFLRLALRTYGRRLAYLPVGRVSKTPASPVYVVOGCPVDAHLVPLEEPP 246
Db 186 LCGMRFTLGTFLRLALRTYGRRLAYLPVGRVSKTPASPVYVVOGCPVDAHLVPLEEPP 244
QY 247 SHMTVVPDEDFVLVLLSHLSGSEMFAPMGRCAGVNHILFYVAGVSRAMLRLFLAM 306
Db 245 SHMTVVPDEDFVLVLLSHLSGSEMFAPMGRCAGVNHILFYVAGVSRAMLRLFLAM 304
QY 307 EKGRIMEYECPLYLVYVVAFRLEPKDKGFAVDGLMVSFAVGOVHPHYFMVSGCV 366
Db 305 OKKGMELDCPLYLVYVVAFRLEPKDKGFAVDGLMVSFAVGOVHPHYFMVSGCV 364
QY 367 EPPPSMKPOQMPPEEP 383
Db 365 DAPSGRDSRRGPPPEEP 381
```

RESULT 9
US-09-796-487-4
Sequence 4, Application US/09796487
Patent No. US20020042358A1
GENERAL INFORMATION:
APPLICANT: Spiegel, Sarah
TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
FILE REFERENCE: 07320001aa (2033957-0001)
CURRENT APPLICATION NUMBER: US/09/796,487
CURRENT FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: US 60/186,532
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: US 09/530,868
PRIOR FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 381
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1)..(381)
OTHER INFORMATION: SEQ ID NO 4 is the peptide sequence of mSPK1a in Figure 3, corresponding to amino acid residue 124 to 504 of SPK1a of GenBank sequence AF068748. SEQ ID NO 4 is equivalent to SEQ ID NO 1 that is the amino acid sequence of SPK1a in Figure 1.
OTHER INFORMATION: Q ID NO 1 that is the amino acid sequence of SPK1a in Figure 1.
AUTHORS: Kohama et al.,
TITLE: Molecular cloning and functional characterization of murine sphingosine kinase
JOURNAL: Journal of Biological Chemistry
VOLUME: 273
ISSUE: 37
PAGES: 23722-23728
DATE: 1998
DATABASE ACCESSION NUMBER: AAC61697
DATABASE ENTRY DATE: 1998-09-26
RELEVANT RESIDUES: (124)..(504)
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: AAC61697

DATABASE ENTRY DATE: 1998-09-26
RELEVANT RESIDUES: (124)..(504)
US-09-796-487-4

```
Query Match 80.7%; Score 1627.5; DB 10; Length 381;  
Best Local Similarity 81.4%; Pred. No. 7,4e-151;  
Matches 307; Conservative 33; Mismatches 36; Indels 1; Gaps 1;  
QY 7 PGCVLPKPCRYLVLLNPGKGRKALQLFRSHVOPLLAEAEISFTLMTERRNNHARELVRS 66  
Db 6 PRLGLRPRCRYLVLLNPGKGRKALQLFRSHVOPFLPEAEITFKILTERKNHARELVCA 65  
QY 67 EELGRMDALVYVSGDGLMHEVYVNGLMERPMETAIQKPLCSIPAGSGNALAASVNHAGY 126  
Db 66 EELGHMDALVYVSGDGLMHEVYVNGLMERPMETAIQKPLCSIPAGSGNALAASVNHAGY 125  
QY 127 EGYTNEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSLACGFIADVDLESEKYYR 186  
Db 126 EGYTNEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSLACGFIADVDLESEKYYR 185  
QY 187 LCGMRFTLGTFLRLALRTYGRRLAYLPVGRVSKTPASPVYVVOGCPVDAHLVPLEEPP 246  
Db 186 LCGMRFTLGTFLRLALRTYGRRLAYLPVGRVSKTPASPVYVVOGCPVDAHLVPLEEPP 244  
QY 247 SHMTVVPDEDFVLVLLSHLSGSEMFAPMGRCAGVNHILFYVAGVSRAMLRLFLAM 306  
Db 245 SHMTVVPDEDFVLVLLSHLSGSEMFAPMGRCAGVNHILFYVAGVSRAMLRLFLAM 304  
QY 307 EKGRIMEYECPLYLVYVVAFRLEPKDKGFAVDGLMVSFAVGOVHPHYFMVSGCV 366  
Db 305 OKKGMELDCPLYLVYVVAFRLEPKDKGFAVDGLMVSFAVGOVHPHYFMVSGCV 364  
QY 367 EPPPSMKPOQMPPEEP 383  
Db 365 DAPSGRDSRRGPPPEEP 381
```

RESULT 10
US-09-796-487-5
Sequence 5, Application US/09796487
Patent No. US20020042358A1
GENERAL INFORMATION:
APPLICANT: Spiegel, Sarah
TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
FILE REFERENCE: 07320001aa (2033957-0001)
CURRENT APPLICATION NUMBER: US/09/796,487
CURRENT FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: US 60/186,532
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: US 09/530,868
PRIOR FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 373
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1)..(373)
OTHER INFORMATION: SEQ ID NO 5 is the peptide sequence of Mouse SPK1a in Fig. 2, corresponding to amino acid residue 131 to 504 of SPK1a of GenBank sequence AF068748. SEQ ID NO 5 is equivalent to SEQ ID NO 1 that is the amino acid sequence of SPK1a in Figure 1.
OTHER INFORMATION: Q ID NO 1 that is the amino acid sequence of SPK1a in Figure 1.
AUTHORS: Kohama et al.,
TITLE: Molecular cloning and functional characterization of murine sphingosine kinase
JOURNAL: Journal of Biological Chemistry
VOLUME: 273
ISSUE: 37
PAGES: 23722-23728
DATE: 1998
DATABASE ACCESSION NUMBER: AAC61697
DATABASE ENTRY DATE: 1998-09-26
RELEVANT RESIDUES: (132)..(504)
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: AAC61697

Query Match 79.8%; Score 1609.5; DB 10; Length 373;
Best Local Similarity 81.3%; Pred. No. 4,1e-149;
Matches 304; Conservative 33; Mismatches 36; Indels 1; Gaps 1;


```

: GENERAL INFORMATION:
: APPLICANT: RASTELLI, LUCA
: TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
: FILE REFERENCE: 10716-08
: CURRENT APPLICATION NUMBER: US/09/784,810A
: CURRENT FILING DATE: 2001-02-14
: PRIOR FILING DATE: 2000-02-14
: PRIOR APPLICATION NUMBER: 60/191,261
: PRIOR FILING DATE: 2000-03-22
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO: 4
: LENGTH: 382
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-784-810A-4

```

```

Query Match      81.3%; Score 1639.5; DB 10; Length 382;
Best Local Similarity 80.9%; Pred. No. 5e-152;
Matches 310; Conservative 33; Mismatches 39; Indels 1; Gaps 1;

```

```

QY 1 MDPAGSPGVLPKPCRYVLINLPNGGKALQLFPSHVOPILAEISFTLMTERRNHA 60
   1 MDPVPCRGGLLPKPCRYVLINLPNGGKALQLFPSHVOPILAEISFTLMTERRNHA 60
DB 1 MDPVPCRGGLLPKPCRYVLINLPNGGKALQLFPSHVOPILAEISFTLMTERRNHA 60
QY 61 RELVASELGRWALVYNSGDLGMEHYNGLMERDWTAIQKPLCSLPGSGNALASL 120
   61 RELVASELGRWALVYNSGDLGMEHYNGLMERDWTAIQKPLCSLPGSGNALASL 120
DB 61 RELVASELGRWALVYNSGDLGMEHYNGLMERDWTAIQKPLCSLPGSGNALASL 120
QY 121 NHYAGYEQVNTNEDLLNCTLLCRRLSPNNLSLHTASGLRFLSVLSLWGFVADVLE 180
   121 NHYAGYEQVNTNEDLLNCTLLCRRLSPNNLSLHTASGLRFLSVLSLWGFVADVLE 180
DB 121 NHYAGYEQVNTNEDLLNCTLLCRRLSPNNLSLHTASGLRFLSVLSLWGFVADVLE 180
QY 181 SEKYRRLGEMRFTLGFTLRLAALRTYRGRLAYLPVGRVSKTPASPVVQGPVDANLVP 240
   181 SEKYRRLGEMRFTLGFTLRLAALRTYRGRLAYLPVGRVSKTPASPVVQGPVDANLVP 240
DB 181 SEKYRRLGEMRFTLGFTLRLAALRTYRGRLAYLPVGRVSKTPASPVVQGPVDANLVP 240
QY 241 LEEPVSHMTVVPDEDFVLYLALSHLSGEMFAAPMGRCAAGVMHLYFYRAGVSRAALL 300
   241 LEEPVSHMTVVPDEDFVLYLALSHLSGEMFAAPMGRCAAGVMHLYFYRAGVSRAALL 300
DB 241 LEEPVSHMTVVPDEDFVLYLALSHLSGEMFAAPMGRCAAGVMHLYFYRAGVSRAALL 300
QY 301 RFLFAMEGRHMEYECPLYVYVPAFRLPEKDGKGFAYDGLMSEAVQGVHPNRYFM 360
   301 RFLFAMEGRHMEYECPLYVYVPAFRLPEKDGKGFAYDGLMSEAVQGVHPNRYFM 360
DB 301 RFLFAMEGRHMEYECPLYVYVPAFRLPEKDGKGFAYDGLMSEAVQGVHPNRYFM 360
QY 361 MYSGCVPEPPSMKPOOMPPPEEP 382
   361 MYSGCVPEPPSMKPOOMPPPEEP 382
DB 361 MYSGCVPEPPSMKPOOMPPPEEP 382

```

```

RESULT 5
US-09-970-516-6
: Sequence 6, Application US/09/70516
: Patent No. US20020099029A1
: GENERAL INFORMATION:
: APPLICANT: No. US20020099029A1artis AG
: TITLE OF INVENTION: Induction of blood vessel formation through administration of
: FILE REFERENCE: 4-31617
: CURRENT APPLICATION NUMBER: US/09/970,516
: CURRENT FILING DATE: 2001-10-04
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO: 6
: LENGTH: 382
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-970-516-6

```

```

Query Match      81.0%; Score 1632.5; DB 10; Length 382;
Best Local Similarity 80.7%; Pred. No. 2.4e-151;

```

```

Matches 309; Conservative 33; Mismatches 40; Indels 1; Gaps 1;
QY 1 MDPAGSPGVLPKPCRYVLINLPNGGKALQLFPSHVOPILAEISFTLMTERRNHA 60
   1 MDPVPCRGGLLPKPCRYVLINLPNGGKALQLFPSHVOPILAEISFTLMTERRNHA 60
DB 1 MDPVPCRGGLLPKPCRYVLINLPNGGKALQLFPSHVOPILAEISFTLMTERRNHA 60
QY 61 RELVASELGRWALVYNSGDLGMEHYNGLMERDWTAIQKPLCSLPGSGNALASL 120
   61 RELVASELGRWALVYNSGDLGMEHYNGLMERDWTAIQKPLCSLPGSGNALASL 120
DB 61 RELVASELGRWALVYNSGDLGMEHYNGLMERDWTAIQKPLCSLPGSGNALASL 120
QY 121 NHYAGYEQVNTNEDLLNCTLLCRRLSPNNLSLHTASGLRFLSVLSLWGFVADVLE 180
   121 NHYAGYEQVNTNEDLLNCTLLCRRLSPNNLSLHTASGLRFLSVLSLWGFVADVLE 180
DB 121 NHYAGYEQVNTNEDLLNCTLLCRRLSPNNLSLHTASGLRFLSVLSLWGFVADVLE 180
QY 181 SEKYRRLGEMRFTLGFTLRLAALRTYRGRLAYLPVGRVSKTPASPVVQGPVDANLVP 240
   181 SEKYRRLGEMRFTLGFTLRLAALRTYRGRLAYLPVGRVSKTPASPVVQGPVDANLVP 240
DB 181 SEKYRRLGEMRFTLGFTLRLAALRTYRGRLAYLPVGRVSKTPASPVVQGPVDANLVP 240
QY 241 LEEPVSHMTVVPDEDFVLYLALSHLSGEMFAAPMGRCAAGVMHLYFYRAGVSRAALL 300
   241 LEEPVSHMTVVPDEDFVLYLALSHLSGEMFAAPMGRCAAGVMHLYFYRAGVSRAALL 300
DB 241 LEEPVSHMTVVPDEDFVLYLALSHLSGEMFAAPMGRCAAGVMHLYFYRAGVSRAALL 300
QY 301 RFLFAMEGRHMEYECPLYVYVPAFRLPEKDGKGFAYDGLMSEAVQGVHPNRYFM 360
   301 RFLFAMEGRHMEYECPLYVYVPAFRLPEKDGKGFAYDGLMSEAVQGVHPNRYFM 360
DB 301 RFLFAMEGRHMEYECPLYVYVPAFRLPEKDGKGFAYDGLMSEAVQGVHPNRYFM 360
QY 361 MYSGCVPEPPSMKPOOMPPPEEP 382
   361 MYSGCVPEPPSMKPOOMPPPEEP 382
DB 361 MYSGCVPEPPSMKPOOMPPPEEP 382

```

```

RESULT 6
US-09-817-676A-15
: Sequence 15, Application US/09/817676A
: Patent No. US20020042101A1
: GENERAL INFORMATION:
: APPLICANT: Spiegell, Sarah
: TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
: FILE REFERENCE: 00170/HG
: CURRENT APPLICATION NUMBER: US/09/817,676A
: CURRENT FILING DATE: 2001-03-26
: PRIOR FILING DATE: 2000-04-03
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 15
: LENGTH: 388
: TYPE: PRT
: ORGANISM: Mus musculus
: PUBLICATION INFORMATION:
: TITLE: Molecular cloning and functional characterization of
: JOURNAL: J. Biol. Chem.
: VOLUME: 273
: ISSUE: 37
: PAGES: 23722-23728
: DATE: 1998-09-11
: DATABASE ACCESSION NUMBER: AAC61698
: DATABASE ENTRY DATE: 1998-09-26
US-09-817-676A-15

```

```

Query Match      80.9%; Score 1630.5; DB 10; Length 388;
Best Local Similarity 81.7%; Pred. No. 3.9e-151;
Matches 308; Conservative 32; Mismatches 36; Indels 1; Gaps 1;

```

```

QY 7 PGVYPRPCRYVLINLPNGGKALQLFPSHVOPILAEISFTLMTERRNHAELVRS 66
   13 PGVYPRPCRYVLINLPNGGKALQLFPSHVOPILAEISFTLMTERRNHAELVRS 66
DB 13 PGVYPRPCRYVLINLPNGGKALQLFPSHVOPILAEISFTLMTERRNHAELVRS 66
QY 67 EELGRWALVYNSGDLGMEHYNGLMERDWTAIQKPLCSLPGSGNALASLHNYAGY 126
   67 EELGRWALVYNSGDLGMEHYNGLMERDWTAIQKPLCSLPGSGNALASLHNYAGY 126

```

Qy 301 RLFLAMEGRHMEYECYLYVYPVAFRLERPKDGGVFAVDGELMSEAVOGVHPNYFW 360
Db 301 RLFLAMEGRHMEYECYLYVYPVAFRLERPKDGGVFAVDGELMSEAVOGVHPNYFW 360
Qy 361 MVSGCVERPPPSWKPOQMPPEEPL 384
Db 361 MVSGCVERPPPSWKPOQMPPEEPL 384

RESULT 2

US-09-796-487-3
; Sequence 3, Application US/09796487
; Patent No. US20020042358A1
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin kinase, cloning, expression and methods of use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(384)
; OTHER INFORMATION: SEQ ID NO 3 is the peptide sequence of hSPHK1 in Fig. 3, correspond-
; OTHER INFORMATION: nding to amino acid residue 1 to 384 of Homo sapiens SPHK-1 of Ge-
; OTHER INFORMATION: nbank sequence Accession Number AAF73423.
; PUBLICATION INFORMATION:
; AUTHORS: Nava et al.
; TITLE: Functional characterization of human sphingosine kinase-1
; JOURNAL: FEBS Lett.
; VOLUME: 473
; ISSUE: 1
; PAGES: 81-84
; DATE: 2000
; DATABASE ACCESSION NUMBER: AAF73423
; DATABASE ENTRY DATE: 2000-06-01
; RELEVANT RESIDUES: (1)..(384)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AAF73423
; DATABASE ENTRY DATE: 2000-06-01
; RELEVANT RESIDUES: (1)..(384)
US-09-796-487-3

Query Match 99.6%; Score 2008; DB 10; Length 384;
Best Local Similarity 99.2%; Pred. No. 4,6e-188;
Matches 381; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDPAGPRGVLPKPCRVLYLNPGRGKKAQQLFRSHQPLLAEEISFTLMLTERRNHA 60
Db 1 MDPAGPRGVLPKPCRVLYLNPGRGKKAQQLFRSHQPLLAEEISFTLMLTERRNHA 60
Qy 61 RELVSEELGRMDALVYVSGDGLMHEVYVNGLMERPDWETAIOKPLCSPAGSGNLAASL 120
Db 61 RELVSEELGRMDALVYVSGDGLMHEVYVNGLMERPDWETAIOKPLCSPAGSGNLAASL 120
Qy 121 NHYAGYEQVNTNEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSLAWGFIADVLE 180
Db 121 NHYAGYEQVNTNEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSLAWGFIADVLE 180
Qy 181 SEKVRLEGMRFTLTGTLRLAALRTYRGLAVLPYGRVSGKTPASPVVVOGPDVAHLVP 240
Db 181 SEKVRLEGMRFTLTGTLRLAALRTYRGLAVLPYGRVSGKTPASPVVVOGPDVAHLVP 240

Qy 241 LEEPVSMTVVPDEDFVLVLLALSHLSEMEFAAPRCACAGVNHLLFYRAGYSRAMLL 300
Db 241 LEEPVSMTVVPDEDFVLVLLALSHLSEMEFAAPRCACAGVNHLLFYRAGYSRAMLL 300
Qy 301 RLFLAMEGRHMEYECYLYVYPVAFRLERPKDGGVFAVDGELMSEAVOGVHPNYFW 360
Db 301 RLFLAMEGRHMEYECYLYVYPVAFRLERPKDGGVFAVDGELMSEAVOGVHPNYFW 360
Qy 361 MVSGCVERPPPSWKPOQMPPEEPL 384
Db 361 MVSGCVERPPPSWKPOQMPPEEPL 384

RESULT 3

US-09-784-810A-2
; Sequence 2, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-810A-2

Query Match 97.3%; Score 1962; DB 10; Length 384;
Best Local Similarity 97.7%; Pred. No. 1,4e-183;
Matches 375; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MDPAGPRGVLPKPCRVLYLNPGRGKKAQQLFRSHQPLLAEEISFTLMLTERRNHA 60
Db 1 MDPAGPRGVLPKPCRVLYLNPGRGKKAQQLFRSHQPLLAEEISFTLMLTERRNHA 60
Qy 61 RELVSEELGRMDALVYVSGDGLMHEVYVNGLMERPDWETAIOKPLCSPAGSGNLAASL 120
Db 61 RELVSEELGRMDALVYVSGDGLMHEVYVNGLMERPDWETAIOKPLCSPAGSGNLAASL 120
Qy 121 NHYAGYEQVNTNEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSLAWGFIADVLE 180
Db 121 NHYAGYEQVNTNEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSLAWGFIADVLE 180
Qy 181 SEKVRLEGMRFTLTGTLRLAALRTYRGLAVLPYGRVSGKTPASPVVVOGPDVAHLVP 240
Db 181 SEKVRLEGMRFTLTGTLRLAALRTYRGLAVLPYGRVSGKTPASPVVVOGPDVAHLVP 240
Qy 241 LEEPVSMTVVPDEDFVLVLLALSHLSEMEFAAPRCACAGVNHLLFYRAGYSRAMLL 300
Db 241 LEEPVSMTVVPDEDFVLVLLALSHLSEMEFAAPRCACAGVNHLLFYRAGYSRAMLL 300
Qy 301 RLFLAMEGRHMEYECYLYVYPVAFRLERPKDGGVFAVDGELMSEAVOGVHPNYFW 360
Db 301 RLFLAMEGRHMEYECYLYVYPVAFRLERPKDGGVFAVDGELMSEAVOGVHPNYFW 360
Qy 361 MVSGCVERPPPSWKPOQMPPEEPL 384
Db 361 MVSGCVERPPPSWKPOQMPPEEPL 384

RESULT 4

US-09-784-810A-4
; Sequence 4, Application US/09784810A
; Patent No. US20020082203A1

Mon May 16 10:50:39 2003

us-09-937-060a-5.rapb

Page 1

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:49:40 ; Search time 23 Seconds
(without alignments)
1536.429 Million cell updates/sec

Title: US-09-937-060a-5
RefSeq score: 2016
Sequence: 1 MDPAGGPGVLPKRCVLYL.....CPEPPSKWKQMPPEPL 384

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2016	100.0	384	10	US-09-970-516-2
2	2008	99.6	384	10	US-09-796-487-3
3	1962	97.3	384	10	US-09-784-810A-2
4	1639.5	81.3	382	10	US-09-784-810A-4
5	1632.5	81.0	382	10	US-09-970-516-6
6	1630.5	80.9	388	10	US-09-817-676A-15
7	1630.5	80.9	388	10	US-09-796-487-2
8	1627.5	80.7	381	10	US-09-796-487-1
9	1627.5	80.7	381	10	US-09-796-487-4
10	1609.5	79.8	373	10	US-09-796-487-5
11	938.5	46.6	204	10	US-09-817-676A-12
12	893	44.3	617	10	US-09-817-676A-14
13	886.5	44.0	618	10	US-09-970-516-4
14	886.5	44.0	618	10	US-09-970-516-6
15	374.5	18.6	415	10	US-09-784-810A-15
16	374	18.6	392	10	US-09-796-487-8
17	356.5	17.7	424	10	US-09-796-487-6
18	356.5	17.7	453	10	US-09-784-810A-12
19	355	17.6	380	10	US-09-784-810A-14

20	345	17.1	403	10	US-09-796-487-7	Sequence 7, Appl 1
21	345	17.1	436	10	US-09-784-810A-13	Sequence 13, Appl 1
22	251	12.5	471	10	US-09-784-810A-6	Sequence 6, Appl 1
23	227.5	11.3	326	10	US-09-784-810A-11	Sequence 11, Appl 1
24	160	7.9	365	9	US-09-738-626-6336	Sequence 29, Appl 1
25	159	7.9	382	10	US-09-784-810A-29	Sequence 29, Appl 1
26	155.5	7.7	421	10	US-09-948-920-50	Sequence 50, Appl 1
27	147	7.3	240	10	US-09-796-487-10	Sequence 10, Appl 1
28	124	6.2	296	10	US-09-815-242-4883	Sequence 4883, Ap
29	124	6.2	362	10	US-09-815-242-10793	Sequence 10793, A
30	111	5.5	294	10	US-09-815-242-13334	Sequence 13334, A
31	106	5.3	315	10	US-09-815-242-12862	Sequence 12862, A
32	100	5.0	294	10	US-09-815-242-5505	Sequence 5505, Ap
33	100	5.0	315	10	US-09-815-242-12494	Sequence 12494, A
34	99.5	4.9	1065	10	US-09-771-161A-239	Sequence 239, App
35	99	4.9	3782	9	US-09-860-846-4	Sequence 4, Appl 1
36	99	4.9	3782	9	US-09-968-384B-4	Sequence 4, Appl 1
37	99	4.9	3782	10	US-09-861-289-4	Sequence 4, Appl 1
38	92	4.6	782	9	US-09-908-193-47	Sequence 47, Appl 1
39	92	4.6	1429	10	US-09-996-617-2	Sequence 2, Appl 1
40	92	4.6	1429	10	US-09-931-071-2	Sequence 2, Appl 1
41	92	4.6	1473	9	US-09-388-221-2	Sequence 4, Appl 1
42	91	4.5	1399	9	US-09-388-221-4	Sequence 6, Appl 1
43	91	4.5	1443	9	US-09-388-221-6	Sequence 3536, Ap
44	89.5	4.4	341	9	US-09-738-626-3536	Sequence 921, App
45	89.5	4.4	974	9	US-10-101-464A-921	

ALIGNMENTS

RESULT 1
US-09-970-516-2
Sequence 2, Application US/0970516
Patent No. US20020099029A1
GENERAL INFORMATION:
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-31617
CURRENT FILING DATE: 2001-10-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
US-09-970-516-2,
Query Match 100.0% Score 2016; DB 10; Length 384;
Best Local Similarity 100.0% Pred 7.5e-189;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MDPAGGPGVLPKRCVLYLNRPGKKAQLQFRSHVQPLAEISTLMTRRNA	60
DB	1	MDPAGGPGVLPKRCVLYLNRPGKKAQLQFRSHVQPLAEISTLMTRRNA	60
QY	61	REIVRSEELKRMALVYVMSGDGIMHEVYVNGIMERPDEWTAIQPLICSLPAGSGNALASL	120
DB	61	REIVRSEELKRMALVYVMSGDGIMHEVYVNGIMERPDEWTAIQPLICSLPAGSGNALASL	120
QY	121	NHYAGYQDVNEDILNCTLLCRRLISPMNLISLHTASQIRLFSVLSLAWCFIADVLE	180
DB	121	NHYAGYQDVNEDILNCTLLCRRLISPMNLISLHTASQIRLFSVLSLAWCFIADVLE	180
QY	181	SEKVRRLGKREFTLGFRLALRTYRGRALVLPVGRVSGKTPASPVVVOQCPVAHILVP	240
DB	181	SEKVRRLGKREFTLGFRLALRTYRGRALVLPVGRVSGKTPASPVVVOQCPVAHILVP	240
QY	241	LEEPVPSHWTVVDEDFVLYLALSHSGSEMPAAMPGRCAACVAMHLFVYRAGVSRAML	300
DB	241	LEEPVPSHWTVVDEDFVLYLALSHSGSEMPAAMPGRCAACVAMHLFVYRAGVSRAML	300

Mon May 12 10:50:39 2003

us-09-937-060a-5.rai

Page 7

[illegible]

Search completed: May 9, 2003, 16:50:25
Job time : 28 secs

APPLICANT: Cosman, David J.
APPLICANT: Anderson, Dirk M.
APPLICANT: Borges, Luis
TITLE OF INVENTION: Family of Immunoregulators Designated Leukocyte Immunoglobulin-
TITLE OF INVENTION: Like Receptors (LIR)
FILE REFERENCE: 2624-A
CURRENT APPLICATION NUMBER: US/09/310,463A
CURRENT FILING DATE: 1999-05-12
EARLIER APPLICATION NUMBER: 08/842,248
EARLIER FILING DATE: 1997-04-24
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
LENGTH: 590
TYPE: PRT
ORGANISM: human
US-09-310-463-22

Query Match 4.3%; Score 86; DB 4; Length 590;
Best Local Similarity 24.8%; Pred. No. 0.86;
Matches 62; Conservative 21; Mismatches 79; Indels 88; Gaps 15;

OY 162 EKYRRLGEMRFL-----GTFRLAALRTYGRGLAYLPVGVGSKTPASP-VVVOQGPV 234
DB 372 QSYRHOAEFSMSPVTSAGGTRCYSAIRSY-----PYL-----LSSPSYPOELVVGSPS 421
OY 235 -DAHLVPL-EEVPVSHMTVPDEDFVLVAL-----LHSHLGSEMFAPMGCAAGVHML 287
DB 422 GPPSLSPGSGTPTPG-----PEDQPLTPGIDPQSLGRHLG----- 458
OY 288 FVVRAGVSRAMLRLFL--AMEKGRHM-EYECPLYLVVPPVVAFRLEPKD----- 333
DB 459 --VVTGVSVAFLVLELLELFLRLRHQSKHRTSAHFYRPAAGAEPKDOGLQKRASPV 516
OY 334 -----GKGVFAVDGELMYSEAVQ-----GVVPRNFMMVSGVEPPSPSK 373
DB 517 ADIOEELINAAYKQPKDGVEMDARAAASEAPQDVYAOLHS--LTLREATEPPPS-- 572
OY 374 PQMPPEEP 383
DB 573 -QERREPAEP 581

RESULT 14
US-08-842-248A-22
Sequence 22, Application US/08842248A
Patent No. 6448035

GENERAL INFORMATION:
APPLICANT: Cosman, David J.
TITLE OF INVENTION: Family of Immunoregulators Designated
TITLE OF INVENTION: Leukocyte Immunoglobulin-Like Receptors (LIR)
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janis C. Henry, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM/PC Compatible
OPERATING SYSTEM: Microsoft Word 7.0
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,248A
FILING DATE: April 24, 1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Henry, Janis C.
REGISTRATION NUMBER: 34,347
REFERENCE/DOCKET NUMBER: 2624
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 590 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-842-248A-22

Query Match 4.3%; Score 86; DB 4; Length 590;
Best Local Similarity 24.8%; Pred. No. 0.86;
Matches 62; Conservative 21; Mismatches 79; Indels 88; Gaps 15;

OY 162 EKYRRLGEMRFL-----GTFRLAALRTYGRGLAYLPVGVGSKTPASP-VVVOQGPV 234
DB 372 QSYRHOAEFSMSPVTSAGGTRCYSAIRSY-----PYL-----LSSPSYPOELVVGSPS 421
OY 235 -DAHLVPL-EEVPVSHMTVPDEDFVLVAL-----LHSHLGSEMFAPMGCAAGVHML 287
DB 422 GPPSLSPGSGTPTPG-----PEDQPLTPGIDPQSLGRHLG----- 458
OY 288 FVVRAGVSRAMLRLFL--AMEKGRHM-EYECPLYLVVPPVVAFRLEPKD----- 333
DB 459 --VVTGVSVAFLVLELLELFLRLRHQSKHRTSAHFYRPAAGAEPKDOGLQKRASPV 516
OY 334 -----GKGVFAVDGELMYSEAVQ-----GVVPRNFMMVSGVEPPSPSK 373
DB 517 ADIOEELINAAYKQPKDGVEMDARAAASEAPQDVYAOLHS--LTLREATEPPPS-- 572
OY 374 PQMPPEEP 383
DB 573 -QERREPAEP 581

RESULT 15
US-09-144-085-2
Sequence 2, Application US/09144085
Patent No. 6280999

GENERAL INFORMATION:
APPLICANT: Gustafsson, Claes
APPLICANT: Betlach, Mary C.
APPLICANT: Ashley, Gary
APPLICANT: Julien, Bryan
APPLICANT: Ziermann, Rainer
TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: 30062-20020,20
CURRENT APPLICATION NUMBER: US/09/144,085
CURRENT FILING DATE: 1998-08-31
EARLIER APPLICATION NUMBER: 09/010,809
EARLIER FILING DATE: 1998-01-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 6095
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-144-085-2

Query Match 4.3%; Score 86; DB 4; Length 6095;
Best Local Similarity 23.4%; Pred. No. 33;
Matches 64; Conservative 29; Mismatches 82; Indels 98; Gaps 12;

OY 1 MDPAGGPRGVLPKPCRYLVILNDRGGKALQDFRSHVPLAEAFISTLMUTERRN-- 58
DB 1427 LDPAG-----TVLV-----TGCTGELGQAIH-----LVRAHGVHVLVTSRGL 1468
OY 59 --HARELYRS-BELGRMDALVWVSGGLMHVYVNGLMERPDMETALOKRCLSPAGSNA 115
DB 1469 APCARELYVSLBELGA-ETVVAACDVSKREYARVLAGID-----AARPL----- 1513

Mon May 1 10:50:39 2003

us-09-937-060a-5.ra1

Page 5

```
Db 11184 GAE-DLDE-----GRSAVLAMVWMOASVJRCDSPEEVPVDPRLFEIGFDSLTVAFDR 11237
OY 208 GHLAVLPVGRVSKTPASPVYVOOGPVDAHNP---LEPVSHHTVYPPDEDFVYALH 265
Db 11238 NRYNRL-----TGLDLP--PVYVFOHTPVALAERISDLEARNNAVAPSD-----H 11283
OY 266 SHLSEMFMAAMGR-----CAQVMHLFVYRAGVSRAMLRLFLAMEKGRHMEY 314
Db 11284 EGAEEKAAAPAGARSGADTGAG-----AGMFLRFRD--AYEDDHYGEF 11326
```

```
RESULT 11
US-09-320-878-18
; Sequence 18, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; EARLIER FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-320-878-18
```

```
Query Match 4.3%: Score 86; DB 3: Length 416;
Best Local Similarity 22.2%: Pred. No. 0.5;
Matches 94; Conservative 51; Mismatches 129; Indels 150; Gaps 23;
```

```
OY 19 VLNPRGKKAQLQFRSHVPL-LEAEISFTLMLTERRNHAR-----EL 63
Db 64 VLADPRSKD-----WRNSTPTTEAEALNHNMLSDPFRHTRKLKVARREFTMRVEL 118
OY 64 VRSEELGRMDALV---VMSGDG---LMHE-----VYNGLMERPDP-----WETAI 101
Db 119 LRPRVOETIVDGLDAMLAADGADLMESELPVITVISELLGVPEPDRAFRWTDAP 178
OY 102 OKPLCSLPAGSGNALASLNHYAGY-----EYVYVND 133
Db 179 VFP--DDPAQAQOTMA---EMSGYLSRLIDSKRGDGEDLSALVRTSDSDGSRITSEE 232
OY 134 LITNCTLLCRRLSPMNLISHTASGLRLEFSVLSLAWGFADVDLESEKRYRLEGMRFT 193
Db 233 LLGMAHILVAGHETVNLIT---ANGM-----YALLSHPD-----264
OY 194 LGTFLRLAALRTYGRGLAYLPVGRVSKTPASPVYVOOGPVDA--HLVPLEPVPSHMTV 251
Db 265 -----QLAALRADMTLLD---GAV-----EEMLRLEGVESATYFEPV-EPVDLDGTV 308
OY 252 VDDDEFVYVYLLSHLSHLSGSEMPFAAPMGRCAGVYMLFVYRAGVSRAMLRLFLAMEKGRH 311
Db 309 IPAGDTVLYV-LADAHKTRPERPDP-----HRDIFRDTAG-----HLAFGHIGH 352
```

```
OY 312 MEYECPLVYVVPVAFR--LEPKDCKGVFAVDGELMYSEAVOQVHNPYFMVSGCVPEP 369
Db 353 FCICAPLARLEARIARALLERCPDLALDVPSELV-----WTFNP--MIRKLALP 402
OY 370 PSKR 373
Db 403 IRWR 406
```

```
RESULT 12
US-09-105-537-39
; Sequence 39, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-39
```

```
Query Match 4.3%: Score 86; DB 4: Length 416;
Best Local Similarity 22.2%: Pred. No. 0.5;
Matches 94; Conservative 51; Mismatches 129; Indels 150; Gaps 23;
```

```
OY 19 VLNPRGKKAQLQFRSHVPL-LEAEISFTLMLTERRNHAR-----EL 63
Db 64 VLADPRSKD-----WRNSTPTTEAEALNHNMLSDPFRHTRKLKVARREFTMRVEL 118
OY 64 VRSEELGRMDALV---VMSGDG---LMHE-----VYNGLMERPDP-----WETAI 101
Db 119 LRPRVOETIVDGLDAMLAADGADLMESELPVITVISELLGVPEPDRAFRWTDAP 178
OY 102 OKPLCSLPAGSGNALASLNHYAGY-----EYVYVND 133
Db 179 VFP--DDPAQAQOTMA---EMSGYLSRLIDSKRGDGEDLSALVRTSDSDGSRITSEE 232
OY 134 LITNCTLLCRRLSPMNLISHTASGLRLEFSVLSLAWGFADVDLESEKRYRLEGMRFT 193
Db 233 LLGMAHILVAGHETVNLIT---ANGM-----YALLSHPD-----264
OY 194 LGTFLRLAALRTYGRGLAYLPVGRVSKTPASPVYVOOGPVDA--HLVPLEPVPSHMTV 251
Db 265 -----QLAALRADMTLLD---GAV-----EEMLRLEGVESATYFEPV-EPVDLDGTV 308
OY 252 VDDDEFVYVYLLSHLSHLSGSEMPFAAPMGRCAGVYMLFVYRAGVSRAMLRLFLAMEKGRH 311
Db 309 IPAGDTVLYV-LADAHKTRPERPDP-----HRDIFRDTAG-----HLAFGHIGH 352
OY 312 MEYECPLVYVVPVAFR--LEPKDCKGVFAVDGELMYSEAVOQVHNPYFMVSGCVPEP 369
Db 353 FCICAPLARLEARIARALLERCPDLALDVPSELV-----WTFNP--MIRKLALP 402
OY 370 PSKR 373
Db 403 IRWR 406
```

```
RESULT 13
US-09-310-463-22
; Sequence 22, Application US/09310463A
; Patent No. 6384203
; GENERAL INFORMATION:
```

OY 186 RLGEKRFITGLRLALRYGRGLAVLPVG--RVGSKTPASPVVQ-----QGPV 234
 Db 853 ALGTPLHDCSLRWLSWVAKAG---YKEPGIARCSSEPMADRLTTPTRHRCQKGPV 909
 OY 235 DAHLVP-----LEPPVPSHMTVVPDEDFVLVLLHSHLGSEMPAARGRA-----AG 283
 Db 910 DINIVAKCNACLSPPCKNKGCTODPVELYRCACPYSYKQDC--TVPIINTIQNFCQHG 968

OY 284 VMHL 287
 Db 969 TCHL 972

RESULT 8

US-09-105-537-37
 ; Sequence 37, Application US/09105537A
 ; Patent No. 6265202

GENERAL INFORMATION:

APPLICANT: Sherman, D.H.

APPLICANT: Liu, H.

APPLICANT: Xue, Y.

APPLICANT: Zhao, L.

TITLE OF INVENTION: DNA encoding methymycin and pikromycin

FILE REFERENCE: 600.438US1

CURRENT APPLICATION NUMBER: US/09/105.537A

CURRENT FILING DATE: 1998-06-26

NUMBER OF SEQ ID NOS: 43

SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 37

LENGTH: 1346

TYPE: PR

ORGANISM: Streptomyces venezuelae

US-09-105-537-37

Query Match

Best Local Similarity 4.3%; Score 86.5; DB 4; Length 1346;
 Best Local Similarity 22.1%; Pred. No. 2.7;

Matches 65; Conservative 17; Mismatches 81; Indels 131; Gaps 14;

OY 81 DGLMEVYVNGLMERPDWETAIOKPLCSL-----PAGSG 113
 Db 854 DGGORHLTTSIAEA--WANGIALDNASILPATGALSPAVPDLPYAFQHRSYWISPAAPG 911
 OY 114 NALAASLNHYAGYEQVTNEDTLTNTLLCRRLSPMNLISHTASGLRFSVLAMGF 173
 Db 912 EAPA-----HTASGREVAETGLAMGP 933
 OY 174 IADVLESEKRYRIGEMRFTIGTFLRLA-----Lr-----TYR 207
 Db 934 GAE-DLDEB---GRSAVILAMVMOAASVLRCDSPREVPDRPLREIGFDSLTAVDPR 987
 OY 208 GRLATLPVGRVSKTPASPVVVOGSPVDAHLVP--LEPPVPSHMTVVPDEDFVLVLLH 265
 Db 988 NRVNML-----TGLQLP--PTVVFQHPPTVALAERISDELAEKNNVAEPSPD-----H 1033
 OY 266 SHLGSEMPAARGRA-----CAAGVMHLFVYRAGVSRAMLRLFLAMEGRHMEY 314
 Db 1034 EQAEBEKAAAPAGANSAGDTGAG-----AGMFRALFRQ--AVEDDRYGEF 1076

RESULT 9

US-09-245-041-15
 ; Sequence 15, Application US/09245041
 ; Patent No. 6274339

GENERAL INFORMATION:

APPLICANT: Moore, K.

APPLICANT: Nagle, D.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT

FILE REFERENCE: 7853-136

CURRENT APPLICATION NUMBER: US/09/245.041

CURRENT FILING DATE: 1999-02-05

EARLIER APPLICATION NUMBER: 60/093,630

EARLIER FILING DATE: 1998-07-21

EARLIER APPLICATION NUMBER: 60/104,978
 ; EARLIER FILING DATE: 1998-10-20
 ; NUMBER OF SEQ ID NOS: 131
 ; SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 15
 LENGTH: 2787

TYPE: PR

ORGANISM: Homo sapiens

US-09-245-041-15

Query Match

Best Local Similarity 4.3%; Score 86.5; DB 4; Length 2787;
 Best Local Similarity 22.4%; Pred. No. 8.5;

Matches 68; Conservative 34; Mismatches 101; Indels 101; Gaps 14;

OY 97 WETAIOKPLCSLPAGSGNALASLNHYAG-----YEOVTNEDTLTNTCLLL 142
 Db 2484 WEAFYLPVOILLFNGYWGGLHFLGRMWCHEHLSHFITSSIHSTSI---CTLTA 2540
 OY 143 CRRLSPMNLISL-----HTASGLRFSVLAMGFADVLESEKRYRIGEMRFTLGT 196
 Db 2541 NARVYTEKILKONPSPHTPV--IPGDLSTLWV----- 2572
 OY 197 FLRLAALTITGR-----LALYLVGRVSKTPASPVVVOGSPVDAHLVPLEE 243
 Db 2573 FLEALILTPQGRORPPGLRACPAVILKCLRPML--APAIAYSRGP--PRDIAPSN 2628
 OY 244 PVPSSHMTVVPDE-----DFVLVLLHSHLGSEMPAARGCAAGVMHLF-- 288
 Db 2629 TLPTRCFERMERKIGTSSFRKSGNGIDPOLNCTKHSK--HSLFRNSSGAVFYL--LFTI 2665
 OY 289 --YVRAGVSRAMT-----LRLFLAMEGRHMEYCPYLV--VYVVARLEPKD 333
 Db 2686 SEFWIOAFVNNYNLASITLIMISKYCKLFHCGDCGCTYIDCNFLFVFLNNSFKQEN 2745
 OY 334 GKGV 337
 Db 2746 GRSV 2749

RESULT 10

US-09-105-537-6
 ; Sequence 6, Application US/09105537A
 ; Patent No. 6265202

GENERAL INFORMATION:

APPLICANT: Sherman, D.H.

APPLICANT: Liu, H.

APPLICANT: Xue, Y.

APPLICANT: Zhao, L.

TITLE OF INVENTION: DNA encoding methymycin and pikromycin

FILE REFERENCE: 600.438US1

CURRENT APPLICATION NUMBER: US/09/105.537A

CURRENT FILING DATE: 1998-06-26

NUMBER OF SEQ ID NOS: 43

SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 6

LENGTH: 11877

TYPE: PR

ORGANISM: Streptomyces venezuelae

US-09-105-537-6

Query Match

Best Local Similarity 4.3%; Score 86.5; DB 4; Length 11877;
 Best Local Similarity 22.1%; Pred. No. 82;

Matches 65; Conservative 17; Mismatches 81; Indels 131; Gaps 14;

OY 81 DGLMEVYVNGLMERPDWETAIOKPLCSL-----PAGSG 113
 Db 11104 DGGORHLTTSIAEA--WANGIALDNASILPATGALSPAVPDLPYAFQHRSYWISPAAPG 11161
 OY 114 NALAASLNHYAGYEQVTNEDTLTNTLLCRRLSPMNLISHTASGLRFSVLAMGF 173
 Db 11162 EAPA-----HTASGREVAETGLAMGP 11183
 OY 174 IADVLESEKRYRIGEMRFTIGTFLRLA-----Lr-----TYR 207

Db 564 TTAELLVORFASIDELAAATIDELALEGVCPTAESIANMFVEDNRRLIEELKEIGVN 623
 QY 241 ---LEEPVSHMTVPDEEVVIALHLH--SHLSEMPAAMPGCACV 284
 Db 624 TQRLPEAPAAESPVRGKFTVLGALPHILTRKAEELIKRAGGVASV 672

RESULT 5
 US-08-936-135-4
 Sequence 4, Application US/08936135
 Patent No. 6054293
 GENERAL INFORMATION:
 APPLICANT: Jessier-Lavigne, Marc
 APPLICANT: Chen, Hang
 APPLICANT: he, Zhigang
 TITLE OF INVENTION: Semaphorin Receptors
 NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:
 ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
 STREET: 75 DENISE DRIVE
 CITY: HILLSBOROUGH
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94010
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 FILING DATE: US/08/936,135
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: OSMAN, RICHARD A
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: UC97-288-2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 343-4341
 TELEFAX: (650) 343-4342
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2584 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Peptide
 US-08-936-135-4

Query Match 4.5%; Score 91; DB 3; Length 2584;
 Best Local Similarity 22.5%; Pred. No. 2.3;
 Matches 67; Conservative 44; Mismatches 123; Indels 64; Gaps 12;
 QY 41 LLAEEISFTMLTERNNHRE---LYRSEELGRDALVMS--GDGLMEVNGIMEEP 95
 Db 2049 LYSERHISVALGTLTHLEARKVALLYSLEHISTYRGINLISPGGLITRASPGL----- 2104
 QY 96 DWETIQPCLSPAGSGNALASLNHYAGTE-----QVTNEDLLTNCITLLICRRLSP 149
 Db 2105 NLEVALTRPMEVTA---LVALGLTHISGLNGLKLASPHISITRPLYSGLDLVARGVALL 2159
 QY 150 NMLLSLHATASGLRFLSVLSLAWGFIADVLESEKYRLG---EMRFTLGP-TRLAAL 203
 Db 2160 LKHISLVSEERLELVSU-----ETVRGLNVALILPEHREGGLGLTLBGL 2204
 QY 204 RTYRGRLA-----YLPVGVGSKTPASPVVVOGQPYDA---HLVLEEPVPSHM 249
 Db 2205 YLISGLVSNIEGLGYGLYLEALVALASPASPILESERILEASNAHNSISLEPRGLNG 2264
 QY 250 TVPPEDEVVLALHSHLSEMEFAPMGRCAAGVNHLYFVRACVSRAMLRLFLAME 307
 Db 2265 LNSPQYS-ALALVYSPRTHRASPLEASP-----LYSLVSNSTHNGLLLELYSTLE 2313

RESULT 6
 US-09-134-001C-5553
 Sequence 5553, Application US/09134001C
 Patent No. 6380370
 GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
 FILE REFERENCE: GTC-007
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/064,964
 PRIOR FILING DATE: 1997-11-08
 PRIOR APPLICATION NUMBER: US 60/055,779
 PRIOR FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 5674
 SEQ ID NO 5553
 LENGTH: 319
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-5553

Query Match 4.3%; Score 87.5; DB 4; Length 319;
 Best Local Similarity 23.7%; Pred. No. 0.22;
 Matches 28; Conservative 26; Mismatches 47; Indels 17; Gaps 4;
 QY 22 NDRGKGAOLDRSHVOPLAE--AESFTMLTERNNHARELYSEELGRDALVMS 79
 Db 35 NGIGEVAKSSLSMCKHLSLQSENKGDIIKCKSIKENTSSDV-----DVLFLIG 85
 QY 80 GDGLMEVYNGIMERPDMETATOKPLCSLPAGSGNALAASLNHYAGTEQVTNEDLLTN 137
 Db 86 GDGTNLELVNGVM-----QYQLNPLGVIPGTFNDFTKTLQLHPN-RTASEQLTTS 137

RESULT 7
 US-09-182-024A-2
 Sequence 2, Application US/09182024A
 Patent No. 6342370
 GENERAL INFORMATION:
 APPLICANT: Connolly, Timothy
 APPLICANT: Rajput, Bhannu
 TITLE OF INVENTION: Human Silt Polypeptide and Polynucleotides Encoding
 FILE REFERENCE: 640100-271
 CURRENT APPLICATION NUMBER: US/09/182,024A
 CURRENT FILING DATE: 1998-10-29
 PRIOR APPLICATION NUMBER: 60/063,946
 PRIOR FILING DATE: 1997-10-31
 PRIOR APPLICATION NUMBER: 60/096,420
 PRIOR FILING DATE: 1998-08-13
 NUMBER OF SEQ ID NOS: 5
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 2
 LENGTH: 1523
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-182-024A-2

Query Match 4.3%; Score 87; DB 4; Length 1523;
 Best Local Similarity 24.2%; Pred. No. 2.9;
 Matches 59; Conservative 26; Mismatches 101; Indels 58; Gaps 10;
 QY 98 ETATQ---KPLCSLPAG-----SGNALASLNHYAGTEQVTNEDLLTNCITLLCR 144
 Db 733 ETVVRCSNKGRLALPGMKPDVTELYLEGNHLTAVPRELSALRHLLTLDLSNNSISMLTN 792
 QY 145 RLSPMLLS-----LHPSAGLRFLSVLSLAWGFIADV-----DLESEKYR 185
 Db 793 YTFSMHSLTLLSYNRLRCIPVHAFNGLSLAVLTLHGNDISSVPESGFNDLSLSHL 852

QY 353 QWBNY 358
Db 285 KLPAHF 290

RESULT 2
US-09-412-545-2
; Sequence 2, Application US/09412545
; Patent No. 6255095
; GENERAL INFORMATION:
; APPLICANT: Prescott, Stephen M.
; APPLICANT: Ding, Li
; APPLICANT: Traet, Elie
; TITLE OF INVENTION: HUMAN DIACYLGLYCEROL KINASE IOTA
; FILE REFERENCE: 1321.2.25
; CURRENT APPLICATION NUMBER: US/09/412.545
; EARLIER FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: 60/103,079
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1065
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-412-545-2

Query Match
Best Local Similarity 24.6%; Score 99.5; DB 4; Length 1065;
Matches 55; Conservative 33; Mismatches 65; Indels 71; Gaps 12;

QY 3 PGGPRGVLPKPCRVLLNPRGKGK--KALQLFRSHVQPLAEALISFTLMTERRNH 59
Db 368 PISPP--LAMP--LVFVNPKSGNGCTKVLQFMWVLPN----- 403

QY 60 ARELYRSEELGRNAL-----VMSGDLMEHVNVGMLERPPWETAIOKPLCSL 108
Db 404 -ROVFDDLSOECPKALLEYRVPNLRLIACGCGDGTVCWILSTLDE--LOLSPPPGYL 459

QY 109 PAGSGNALAASLNHYAGYEQVTNEDLTNCTLLCR--RLISPMLLSLHTASG----- 160
Db 460 PLGTGNDLARTLNNGGCT--TDEPV-----SKLLQVEDGTIVVOLDRRNLHVERNPDLPP 512

QY 161 -----LRLE-SVLSLWAGFIADVDESEKRYRLGEMRF 192
Db 513 EELEDGVCKLEPLWENNYFSL--GDAHVTLLEPHSREANPEKF 554

RESULT 3
US-09-105-537-4
; Sequence 4, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105.537A
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3782
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-4

Query Match
Best Local Similarity 24.1%; Score 99; DB 4; Length 3782;
Matches 103; Conservative 48; Mismatches 189; Indels 88; Gaps 19;

QY 3 PGGPRGVLPKPCRVLLNPRGKGKALQLFRSHVQPLAEALISFTLMTERRNH 61
Db 132 PEGORAVVEGIRHRETLLEGAPDPSASAYAFELLCFVRPAVTAANAAYLVGPADRRADFA 191
QY 62 ELVR-----SEELGRNAL-VMSGDLMEHVNVGMLERPPWETAIOKPLCSL 113
Db 192 DLRLRLPSDLSLAPDSLRVRAADGALAEALTLADSDSPGALSLAGVTAAYOLTG 251
QY 114 NALAASLNHYAGYEQVTNEDLTNCTLLCRLLSPML-----LSLHTASGLR----- 163
Db 252 NAVLALLAHPRQMBELCDREGLAAVAEBETLRYPPODLARVYRGETELAGRLPACAH 311
QY 164 FSVLSLWAGFIADVDESEKX--RUGEMRFTL-----GTPLRL--AALRTYRG 208
Db 312 VVVLTAATGRPEVFTDPERDLARPDAAHILHPACQPVASVLYQLQEVALLRTLAG 371
QY 209 RLATL-----PVGVSKTASP--VVQGGPDAHLVPLEEVPVSHWYVP 253
Db 372 RFPGLRQAGDVLRRRAVAGRPVSSSMRVLLTSFAHHTHYGCL--VPLAMALLA 428
QY 254 DEDVLYLV--ALLSHLSEMFAPMRCAGVMHLEFVR-AGVSRAMLRLFLAMEKG 309
Db 429 AGHEVYASQPALDITDTTSGSLAAYPVG--TDILIHRYVMAGEPRP----- 474
QY 310 RHMEYECPLYVVPYVAF--RLERKDGKGVFAVDGELMVSEAVOGGVHPYFMVWVSGCV 366
Db 475 NH-----PAIARDEARPELDWDHALGIEAIL--APYFHLNANDSVDDLV 519
QY 367 EPPSWKP 374
Db 520 DFARSWOP 527

RESULT 4
US-09-708-426-9
; Sequence 9, Application US/09708426
; Patent No. 6444429
; GENERAL INFORMATION:
; APPLICANT: HAN, YE-SUN
; APPLICANT: YU, YEON-GYU
; APPLICANT: LIM, JAE-HWAN
; TITLE OF INVENTION: GENE CODING FOR DNA LIGASE OF HYPERTHERMOPHILIC BACTERIA AQUT
; FILE REFERENCE: 199579050
; CURRENT APPLICATION NUMBER: US/09/708.426
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: KR99-49591
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 9
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Rhodothermus marinus
US-09-708-426-9

Query Match
Best Local Similarity 21.8%; Score 93; DB 4; Length 712;
Matches 63; Conservative 37; Mismatches 109; Indels 80; Gaps 11;

QY 59 HARELYRSEELGRMDALVMSGDLMEHVNGIM-----ERPDMETAIOKPLCS----- 107
Db 401 HNEDEVRSRDIRIGLVVYIRAGDVYIPQVVRPVVEAFRTGNERP--WRMPERCSCGSQLVR 459
QY 108 LPAGSGNALAAS-----LNHYAGYEQVTNEDLTNCTLLCR--LLSPM----- 150
Db 460 LPGEADYVCVADCPACQFVRLLEHFAGRDAMDIEGMSQVVAROLAESGLVRLSDLYRLK 519
QY 151 --NLSLHTASGLRLFSVLSLWAGFIADVDESEKRYRLGEMRFITGFLRLAALRTYRG 208
Db 520 LEDLKLKGFAPFETRRNRLLRA-----IEASQORLSRLRLFLG-----IRHKG 563
QY 209 RLATYLPVGRVSGKTPASPVVVOO-----GPVDAHLVP----- 240

Mon May 1 10:50:39 2003

us-09-937-060a-5.ra1

Page 1

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: May 9, 2003, 16:46:18 ; Search time 15 Seconds
(without alignments)
753.227 Million cell updates/sec

Title: US-09-937-060A-5
Perfect score: 2016
Sequence: 1 MDPACGPGVLPKRCRVLV.....CVEPPSMKPOQMPPEEPL 384

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2.6/ptodata/1/1aa/5a_COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5b_COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/6a_COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/6b_COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/6c_COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/bckfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	5.1	324	US-09-134-001C-4774	Sequence 4774, App1
2	99.5	4.9	1065	US-09-412-545-2	Sequence 2, App1
3	99	4.9	3782	US-09-105-537-4	Sequence 4, App1
4	93	4.6	712	US-09-708-426-9	Sequence 9, App1
5	91	4.5	2584	US-08-936-135-4	Sequence 4, App1
6	87.5	4.3	319	US-09-134-001C-5553	Sequence 5553, App1
7	87	4.3	1523	US-09-182-0248-2	Sequence 2, App1
8	86.5	4.3	1346	US-09-105-537-37	Sequence 3, App1
9	86.5	4.3	2787	US-09-245-041-15	Sequence 15, App1
10	86.5	4.3	11877	US-09-320-878-18	Sequence 18, App1
11	86	4.3	416	US-09-105-537-6	Sequence 6, App1
12	86	4.3	416	US-09-105-537-18	Sequence 18, App1
13	86	4.3	416	US-09-105-537-39	Sequence 39, App1
14	86	4.3	590	US-08-310-463-22	Sequence 22, App1
15	86	4.3	6095	US-08-842-248A-22	Sequence 22, App1
16	85.5	4.2	1580	US-08-144-085-2	Sequence 11, App1
17	85.5	4.2	1580	US-08-804-196-5	Sequence 11, App1
18	84.5	4.2	471	US-08-176-427B-11	Sequence 11, App1
19	84.5	4.2	471	US-08-356-060A-34	Sequence 34, App1
20	84.5	4.2	471	US-08-460-900C-34	Sequence 34, App1
21	84.5	4.2	471	US-08-757-230X-8	Sequence 8, App1
22	84.5	4.2	471	US-08-674-509B-34	Sequence 34, App1
23	84.5	4.2	471	US-08-954-698-34	Sequence 34, App1
24	84.5	4.2	471	US-08-957-874-34	Sequence 34, App1
25	84	4.2	423	US-08-760-797A-1	Sequence 1, App1
26	84	4.2	424	US-08-932-929B-1	Sequence 1, App1
27	83.5	4.1	1346	US-09-320-878-4	Sequence 4, App1

28	83	4.1	424	US-08-760-797A-3	Sequence 3, App1
29	83	4.1	424	US-08-932-929B-3	Sequence 3, App1
30	83	4.1	435	US-08-031-538-11	Sequence 11, App1
31	83	4.1	520	PCT-US95-04801-5	Sequence 5, App1
32	82.5	4.1	471	US-09-325-256-26	Sequence 26, App1
33	82	4.1	840	US-08-500-857A-6	Sequence 6, App1
34	80.5	4.0	287	US-07-952-817-24	Sequence 24, App1
35	80.5	4.0	287	US-08-402-467B-2	Sequence 2, App1
36	80	4.0	1447	US-08-340-406-19	Sequence 19, App1
37	80	4.0	1447	US-08-656-055-19	Sequence 19, App1
38	80	4.0	1447	US-08-954-668-19	Sequence 19, App1
39	80	4.0	1447	US-09-268-140-5	Sequence 5, App1
40	80	4.0	1447	US-08-918-658-19	Sequence 19, App1
41	80	4.0	1447	PCT-US95-13233-19	Sequence 19, App1
42	79	3.9	928	US-08-841-483-4	Sequence 4, App1
43	79	3.9	928	US-09-382-911-4	Sequence 4, App1
44	79	3.9	1117	US-08-841-483-6	Sequence 6, App1
45	79	3.9	1117	US-09-382-911-6	Sequence 6, App1

ALIGNMENTS

RESULT 1
US-09-134-001C-4774
Sequence 4774, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Dorelle-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134/001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4774
LENGTH: 324
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4774

Query Match 5.1% Score 102: DB 4: Length 324:
Best Local Similarity 19.7%: Pred No 0.0052:
Matches 72: Conservative 58: Mismatches 126: Indels 110: Gaps 18:

QY	16	RIVVLNPGSGKGLDIFRSHVOPLAFAEISFTMLTERNNARELYRSEELG-----	70
DB	12	RAVILNPGSGKGLDIFRSHVOPLAFAEISFTMLTERNNARELYRSEELG-----	70
QY	71	-----RVDALVMSGGLMEHYVNGIMERPMEYAIQKPLCSLPAGSNAALASLNHY	123
DB	59	AERALESEYDILLIAGGGGTENVVNGIAEOPN-----RPKLGVIIPKGTVDNDFGRALH--	111
QY	124	AGYEYVNDLITNCTLLCRRLSLPNN-LISLHTASGURLPSVLSLAWGFIADVDESE	182
DB	112	-----LPSDINGAIDVILIDHTTK-----VIGGMNN	138
QY	183	KYRRLGEMFTGLFLRLAALRYKGLAYLPGRVSKTPASPVVVOGPDVAHLPLE	242
DB	139	RY-----FINLAA-----GKLT-----QVSEYTP-SKLKSIVGPF-AVYIKGF	175
QY	243	EPVPSHWTV-----VPDEDFVVLALHSHLSGSMFAAPGRCACAVMLFVYRA	292
DB	176	EMLPQKKAADVIEYDDVIFGEALLRLGLTSMNGFEKL-VPAKIDDDYFLILYK	234
QY	293	GVSRAALLRLFLAMEKSGHMEYECPIYLVYVPAFLPKDKGVFAVDSGLAMSEAVOG	352
DB	235	ANLALGHTMTLASRGHTKH--PKVITAKAKSINISS-----FTDMOLNVGDTYGG	284

Matches 307; Conservative 33; Mismatches 36; Indels 1; Gaps 1;

```

OY 7 PRGVLPKPCRVLYLNPFGKGAQLOLFRSHVOPPLAEAEISFTLM.TERRNHARELYRS 66
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 6 PRGLLPKPCRVLYLNPFGKGAQLOLFRSHVOPPLAEAEITFKLITERKNHARELYCA 65
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 67 BELGRWDALVYNSGDLMEHEVYNGIMERPDMETAIORPICSLPAGSGNALAASLNHYAGY 126
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 66 BELGHWDAIAMSQDGLMEHEVYNGIMERPDMETAIORPICSLPAGSGNALAASLNHYAGY 125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 127 EOYINEDLLTNCITLLCRRLSPMNLISLHTASGLRLEFSVLSLAWGFIADVOLESEKYYR 186
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 EOYINEDLLTNCITLLCRRLSPMNLISLHTASGLRLEFSVLSLAWGFIADVOLESEKYYR 185
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 187 LGEMFETLGTFLRLAALRYRGLAYLPVGRVSKTPASPVYVOOGFVDAAHLVPLEEPVP 246
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 LGELRFYVGFEEFLASRLTYOGOLATLPVGVASKRPAS-TLVOKGPDVTHLVPLEEPVP 244
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 247 SHMTVVPDEDFVLVALLSHLSGSEMFAPMGRCAGVYMHLEFYVRAGVSRAMLLRLFLAM 306
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 245 SHMTVVPDEDFVLVALLSHLSGSEMFAPMGRCAGVYMHLEFYVRAGVSRAMLLRLFLAM 304
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 307 EKGRHMEYECFYLYVYVYVAFRLPEPKDGVEAFVADGELMSEAVOGVHNPYFMVYSGCV 366
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 305 OKGKHMEIDCPYLVHYVYVAFRLPEPKDGVEAFVADGELMSEAVOGVHNPYFMVYSGCV 364
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 367 EPPPSWKPOOKMPPEEP 383
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 365 DAPSGRDSRRGPPPEEP 381
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Search completed: May 9, 2003, 16:54:55
 Job time : 49 secs

XX	AA556053	standard; Protein: 381 AA.
XX	AA556053	
AC	AA556053	
XX	28-MAR-2000	(first entry)
XX	Mouse sphingosine kinase 1a protein.	
XX	Anticancer: antiproliferative; antiatherosclerotic; stroke; restenosis;	
KW	antineurodegeneration; sphingosine kinase; cell proliferation; apoptosis;	
KW	cancer; diabetic neuropathy; Alzheimer's disease; atherosclerosis.	
XX	Mus musculus.	
OS	Mus musculus.	
PN	MO9961581-A2.	
XX	02-DEC-1999.	
PD	02-DEC-1999.	
XX	25-MAY-1999;	99WO-US11521.
PF	25-MAY-1999;	99WO-US11521.
XX	26-MAY-1998;	98US-0086657.
PR	11-AUG-1998;	98US-0096049.
XX	(DEAN-) OFFICE DEAN RES & GRADUATE EDUCATION.	
PA	(DEAN-) OFFICE DEAN RES & GRADUATE EDUCATION.	
XX	Splegel S;	
PI	Splegel S;	
XX	WPI: 2000-072612/06.	
DR	N-PSDB; AAZ47167.	
XX	New sphingosine kinase, used to treat diseases involving abnormal cell	
PT	proliferation, e.g. cancer -	
XX	Example 1; Fig 1; 116pp; English.	
PS	Example 1; Fig 1; 116pp; English.	
XX	This sequence represents the mouse sphingosine kinase 1a. The coding	
CC	sequence was isolated by searching a dbEST (expressed sequence tag	
CC	database) using amino acid sequence results from sequenced tryptic	
CC	digests of the rat sphingosine kinase. Expression of sphingosine kinase	
CC	in cells results in formation of sphingosine-1-phosphate, a known second	
CC	messenger, and confers serum-independent growth; increases proliferation,	
CC	and suppresses serum-deprivation or ceramide-induced apoptosis. The	
CC	sphingosine kinase nucleic acid is used: (a) to increase sphingosine	
CC	kinase content of cells, specifically for reducing cell death and/or	
CC	are used to screen for agents that inhibit or promote sphingosine kinase	
CC	activity. Agents that reduce sphingosine kinase activity or expression	
CC	are used: (1) to reduce cell proliferation, specifically for treating	
CC	cancer, and (11) to treat diseases associated with abnormal cell	
CC	migration or motility, particularly cancer, restenosis or diabetic	
CC	neuropathy (but also atherosclerosis, stroke and Alzheimer's disease),	
CC	whereas agents that stimulate sphingosine kinase can be used to treat	
CC	conditions associated with reduced cell proliferation, e.g. developmental	
CC	retardation.	
XX		
XX	Sequence 381 AA:	
XX		
XX	Query Match 80.7%; Score 1627.5; DB 21; Length 381;	
XX	Best Local Similarity 81.4%; Pred. No. 1.9e-167;	
XX	Matches 307; Conservative 33; Mismatches 36; Indels 1; Gaps 1;	
QY	7 PGCVLPKPCRYLVLLNPGKGGKALQLFRESHVQPLLAELISFTMLTERNHARELVRS 66	
DB	6 PGGLLPKPCRYLVLLNPGGKGAQLQFOSRVQPLEAEIIFKILITERKHNAREIVCA 65	
QY	67 EELGWMDLAVYVSGGLMHEVYVNGMERPDMTALOKPLCSLPAGSGNLAASLNIHYAY 126	
DB	66 EELGWMDLAVYVSGGLMHEVYVNGMERPDMTALOKPLCSLPAGSGNLAASVNIHYAY 125	

QY	127	EOVNNEDLLTNCCTLLCRLRLSPNNLSLTHASGLRFSVSLAMGFIAVDLDESEKYR	186
Db	126	EOVNNEDLLICTLLTLLCRRRLSPNNLSLTHASGLRFSVSLAMGFIAVDLDESEKYR	185
QY	187	LGEMKFTLGTFLRLAALRYLRCGLAYLPVGVHVGSGKTPASPVVVYQGGVDAHLVLEEPV	246
Db	186	LGELRFTVGTGFERRFLASRLRYVQGLAYLPVGVVASKRPAS- ¹ TLVKGQGVDTNVLPLEEPV	244
QY	247	SHMTVVPPEDEVLVLTALHSLHLSGEMFAAPMGRCACAVMHLFVVRGVSAMLLRLFLAM	306
Db	245	SHMTVVPPEDEVLVLTALHSLHLSGEMFAAPMGRCACAVMHLFVVRGVSAMLLRLFLAM	304
QY	307	EKGRRHMECEPYLVYVVPVVAERLDPKDGKGVFAVYDGLMWSAQAQGVHNPFTMWVSGCV	366
Db	305	OKGKHMELDCPYLVHVPVVAERLEPRSGQGVSYDGLMCAEAQGVHNPFTMWVSGSR	364
QY	367	EPSPSKPOQMPPEEP	383
Db	365	DAPSGRDSRRGPPEEP	381

RESULT 15		
AB080807		
ID	AB080807 standard; protein; 381 AA.	
XX		
AC	AB080807;	
XX		
DT	10-SEP-2002 (first entry)	
XX		
DE	Murine sphingosine kinase 1a (msPK1a) protein sequence.	
XX		
XX	Sphingosine kinase; SPHK; SPHK1a; cytosolic; vasotrophic; antidiabetic;	
KW	neuroprotective; mouse; enzyme.	
XX		
OS	Mus musculus.	
XX		
PN	US2002042358-A1.	
XX		
PD	11-APR-2002.	
XX		
PF	02-MAR-2001; 2001US-0796487.	
XX		
PR	02-MAR-2000; 2000US-186352P.	
XX		
PA	(SPIE/) SPIEGEL S.	
PI	Spiegel S;	
XX		
DR	WPI: 2002-478846/51.	
XX		
PT	New isolated sphingosine kinase, useful in identifying modulators for	
PT	treating e.g. cancer, also related nucleic acid, vectors and	
PT	transformed cells -	
XX		
PS	Disclosure: Fig 1; 2app; English.	
XX		
CC	The invention relates to an isolated sphingosine kinase (SPHK) DNA. Cells	
CC	transformed with SPHK DNA are used to screen for agents that reduce,	
CC	eliminate or promote SPHK activity. Agents that inhibit activity are	
CC	useful for decreasing cell proliferation, e.g. for treating cancer, and	
CC	for treating diseases associated with abnormal migration and motility of	
CC	cells, e.g. retinoblastoma or diabetic neuropathy. Agents that increase	
CC	activity are used to reduce cell death. Antibodies raised against SPHK,	
CC	and primers or oligonucleotides derived from the DNA are useful for	
CC	diagnosis. The antibodies are also useful as therapeutic inhibitors. The	
CC	present sequence represents a murine sphingosine kinase 1a (msPK1a)	
CC	corresponding to residues 124 to 504 of SPHK1a in GenBank accession	
CC	no. AAC61697.	
XX		
XX		
Sequence	381 AA;	
Query Match	80.7%; Score 1627.5; DB 23; Length 381;	
Best Local Similarity	81.4%; Pred NO. 1 9e-167;	

PI Spiegel S;
XX
DR WPI: 2000-072612/06.
XX N-PSDB: AA247168.
XX
XX
XX New sphingosine kinase, used to treat diseases involving abnormal cell
XX proliferation, e.g. cancer -
XX
XX Example 1: Fig 1: 116pp: English.
XX
XX This sequence represents the mouse sphingosine kinase 1a. The coding
XX sequence was isolated by searching a dbEST (expressed sequence tag
XX database) using amino acid sequence results from sequenced tryptic
XX digests of the rat sphingosine kinase. Expression of sphingosine kinase
XX in cells results in formation of sphingosine-1-phosphate, a known second
XX messenger, and confers serum-independent growth; increases proliferation,
XX and suppresses serum-deprivation or ceramide-induced apoptosis. The
XX sphingosine kinase nucleic acid is used: (a) to increase sphingosine
XX kinase content of cells, specifically for reducing cell death and/or
XX increasing cell proliferation; and (b) to produce transfected cells that
XX are used to screen for agents that inhibit or promote sphingosine kinase
XX activity. Agents that reduce sphingosine kinase activity or expression
XX are used: (i) to reduce cell proliferation, specifically for treating
XX cancer, and (ii) to treat diseases associated with abnormal cell
XX migration or motility, particularly cancer, restenosis or diabetic
XX neuropathy (but also atherosclerosis, stroke and Alzheimer's disease),
XX whereas agents that stimulate sphingosine kinase can be used to treat
XX conditions associated with reduced cell proliferation, e.g. developmental
XX retardation.
XX
XX Sequence 388 AA:
SQ
Query Match 80.9%; Score 1630.5; DB 21; Length 388;
Best Local Similarity 81.7%; Pred. No. 9,4e-168;
Matches 308; Conservative 32; Mismatches 36; Indels 1; Gaps 1;
QY 7 PGCVLPKRCVLYLVLPNPGCKGKALQLFQSHVOPPLAEATISFTLMTERRNHARELYRS 66
DB 13 PGGLLPKRCVLYLVLPNPGCKGKALQLFQSHVOPPLAEATISFTLMTERRNHARELYRS 72
QY 67 EELGMDALVMSGDLMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASLVNHVAGY 126
DB 73 EELGMDALVMSGDLMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASLVNHVAGY 132
QY 127 EOVNEDLLNCTLLCRRLSPMNLSTHTASGLRFSVLSLWGFIAVDLESEKYYR 186
DB 133 EOVNEDLLNCTLLCRRLSPMNLSTHTASGLRFSVLSLWGFIAVDLESEKYYR 192
QY 187 LGEHFTLGTFLRLAALRTYRGLAYLPVGRVSKTPASPVYVQGVADVHLPLEDEVP 246
DB 193 LGEHFTLGTFLRLAALRTYRGLAYLPVGRVSKTPASPVYVQGVADVHLPLEDEVP 251
QY 247 SHMTVVPDEDFVLYLVALLSHLSEMEFAAPMGRCACVMIHLYRACVSRAMILRLFLAM 306
DB 252 SHMTVVPDEDFVLYLVALLSHLSEMEFAAPMGRCACVMIHLYRACVSRAMILRLFLAM 311
QY 307 EKGHMEYECPTLYLVYVVAFAFLRPEKDGKGFVAVDGLMSEAVQGVHPNFMVWVSCV 366
DB 312 OKGHMELDCPTLYLVYVVAFAFLRPEKDGKGFVAVDGLMSEAVQGVHPNFMVWVSCV 371
QY 367 EPPSMKPOQMPPEEP 383
DB 372 DAFSGRDSRGPPPEEP 388

DE Murine sphingosine kinase 1b (msPHK1b) protein sequence.
XX
XX Sphingosine kinase; SPHK; SPHK1b; cytosolic; vasotropic; antidiabetic;
XX neuropeptide; mouse; enzyme.
XX
XX Mus musculus.
XX
XX US2002042358-A1.
XX
XX 11-Apr-2002.
XX
XX 02-MAR-2001; 2001US-0796487.
XX
XX 02-MAR-2000; 2000US-186352P.
XX
XX (SPIE/) SPIEGEL S.
XX
XX Spiegel S;
XX
XX WPI: 2002-478846/51.
XX
XX
XX New isolated sphingosine kinase, useful in identifying modulators for
XX treating e.g. cancer, also related nucleic acid, vectors and
XX transformed cells -
XX
XX Disclosure; Fig 1: 24pp: English.
XX
XX The invention relates to an isolated sphingosine kinase (SPHK) DNA. Cells
XX transformed with SPHK DNA are used to screen for agents that reduce,
XX eliminate or promote SPHK activity. Agents that inhibit activity are
XX useful for decreasing cell proliferation, e.g. for treating cancer, and
XX for treating diseases associated with abnormal migration and motility of
XX cells, e.g. restenosis or diabetic neuropathy. Agents that increase
XX activity are used to reduce cell death. Antibodies raised against SPHK,
XX and primers or oligonucleotides derived from the DNA are useful for
XX diagnosis. The antibodies are also useful as therapeutic inhibitors. The
XX present sequence represents a murine sphingosine kinase 1b (msPHK1b)
XX corresponding to residues 1 to 388 of SPHK1b in Genbank Accession
XX no. AAC61698.
XX
XX Sequence 388 AA:
SQ
Query Match 80.9%; Score 1630.5; DB 23; Length 388;
Best Local Similarity 81.7%; Pred. No. 9,4e-168;
Matches 308; Conservative 32; Mismatches 36; Indels 1; Gaps 1;
QY 7 PGCVLPKRCVLYLVLPNPGCKGKALQLFQSHVOPPLAEATISFTLMTERRNHARELYRS 66
DB 13 PGGLLPKRCVLYLVLPNPGCKGKALQLFQSHVOPPLAEATISFTLMTERRNHARELYRS 72
QY 67 EELGMDALVMSGDLMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASLVNHVAGY 126
DB 73 EELGMDALVMSGDLMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASLVNHVAGY 132
QY 127 EOVNEDLLNCTLLCRRLSPMNLSTHTASGLRFSVLSLWGFIAVDLESEKYYR 186
DB 133 EOVNEDLLNCTLLCRRLSPMNLSTHTASGLRFSVLSLWGFIAVDLESEKYYR 192
QY 187 LGEHFTLGTFLRLAALRTYRGLAYLPVGRVSKTPASPVYVQGVADVHLPLEDEVP 246
DB 193 LGEHFTLGTFLRLAALRTYRGLAYLPVGRVSKTPASPVYVQGVADVHLPLEDEVP 251
QY 247 SHMTVVPDEDFVLYLVALLSHLSEMEFAAPMGRCACVMIHLYRACVSRAMILRLFLAM 306
DB 252 SHMTVVPDEDFVLYLVALLSHLSEMEFAAPMGRCACVMIHLYRACVSRAMILRLFLAM 311
QY 307 EKGHMEYECPTLYLVYVVAFAFLRPEKDGKGFVAVDGLMSEAVQGVHPNFMVWVSCV 366
DB 312 OKGHMELDCPTLYLVYVVAFAFLRPEKDGKGFVAVDGLMSEAVQGVHPNFMVWVSCV 371
QY 367 EPPSMKPOQMPPEEP 383
DB 372 DAFSGRDSRGPPPEEP 388

SQ Sequence 333 AA;
 Query Match 86.9%; Score 1752; DB 23; Length 333;
 Best Local Similarity 100.0%; Pred. No. 5e-101;
 Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 MLETRRNHAEELVSEELGKWDALVYVSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAG 111
 Db 1 MLETRRNHAEELVSEELGKWDALVYVSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAG 60

QY 112 SGNALAASLNHYAGYEDVNTNEDLLTNCILLCRRLSPMNLISLHTASGLRLFSVLSLAW 171
 Db 61 SGNALAASLNHYAGYEDVNTNEDLLTNCILLCRRLSPMNLISLHTASGLRLFSVLSLAW 120

QY 172 GFIDVDVLESKRYRLGEMRFTLCTFLIALRTYRGRLAVLPGRGSKTPASPVVVOQ 231
 Db 121 GFIDVDVLESKRYRLGEMRFTLCTFLIALRTYRGRLAVLPGRGSKTPASPVVVOQ 180

QY 232 GPVDAHLVPLEEPYPVSHMTVPDEDFVLVLAHLHSHLGSEMFAPMGRCAGVNHLEFVVR 291
 Db 181 GPVDAHLVPLEEPYPVSHMTVPDEDFVLVLAHLHSHLGSEMFAPMGRCAGVNHLEFVVR 240

QY 292 AGVSRAMLRLFLAMEGRHMEYECPLYVYPVYAFRLPEKDGKGVRAVDGELMAYSEAVQ 351
 Db 241 AGVSRAMLRLFLAMEGRHMEYECPLYVYPVYAFRLPEKDGKGVRAVDGELMAYSEAVQ 300

QY 352 GOVHPNFMVYSGVEPPPSKPOOMPPEEPL 384
 Db 301 GOVHPNFMVYSGVEPPPSKPOOMPPEEPL 333

RESULT 11
 ID AAE07883 standard; Protein: 382 AA.
 AC AAE07883;
 DT 01-NOV-2001 (first entry)
 DE Mouse consensus sphingosine kinase (Sphk) protein.
 KW Mouse; sphingosine kinase; Sphk; restenosis; ischaemia; gene therapy; antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis; cytostatic; tumor necrosis factor-alpha; TNF; atherosclerosis; lymphoma; leukaemia; vasotropic; cell proliferative disorder; vascular disease.
 OS Mus musculus.
 PN WO200160990-A2.
 PD 23-AUG-2001.
 PF 14-FEB-2001; 2001WO-US04789.
 PR 14-FEB-2000; 2000US-0182360.
 PR 22-MAR-2000; 2000US-0191261.
 PA (CUBA-) CURAGEN CORP.
 PA (GETH) GENENTECH INC.
 PI Rastelli L;
 DR WPI: 2001-514770/56.
 DR N-PSDB: AAD14425.
 PT An isolated sphingosine kinase polypeptide useful for treating a
 PT Sphk-associated disorder especially cancer, restenosis or ischemia in a
 PT human -
 PS Example 1: Page 92-94; 107pp; English.
 CC The present invention relates to sphingosine kinase (sphk) polypeptides
 CC and nucleic acids encoding them. sphk is useful for treating a sphk-

CC associated disorder especially cancers such as leukaemia, lymphoma,
 CC ovarian, breast, lung, colon, testicular, stomach and skin,
 CC atherosclerosis, restenosis or ischaemia and cell proliferative disease
 CC or disorder associated with vascular diseases. Sphk gene is used in gene
 CC therapy and antisense-therapy. Sphingolipids serving as signalling
 CC molecules, have recently emerged as regulators of cell growth.
 CC differentiation, diverse cell phenotypes and cell death. Activation of
 CC Sphk by tumor necrosis factor (TNF)-alpha inhibits apoptosis in human
 CC endothelial cells. The present sequence is consensus mouse sphingosine
 CC kinase (Sphk) protein.

SQ Sequence 382 AA;
 Query Match 81.3%; Score 1639.5; DB 22; Length 382;
 Best Local Similarity 80.9%; Pred. No. 9.7e-169;
 Matches 310; Conservative 33; Mismatches 39; Indels 1; Gaps 1;

QY 1 MDPAGGREGVLPRECRVLYLNPFGGKGAQLQFRSHVQPLLAETISFTLMLTERRNA 60
 Db 1 MEPEECPRGLPRPCRVLYLNPFGGKGAQLQFRSHVQPLLAETISFTLMLTERRNA 60

QY 61 RELVSEELGKWDALVYVSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAGSNALASL 120
 Db 61 RELVCAEELGKWDALVYVSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAGSNALASV 120

QY 121 NHYAGYEDVNTNEDLLTNCILLCRRLSPMNLISLHTASGLRLFSVLSLAWGFIADVLE 180
 Db 121 NHYAGYEDVNTNEDLLTNCILLCRRLSPMNLISLHTASGLRLFSVLSLAWGFIADVLE 180

QY 181 SEKYRRLEGMRFTLCTFLIALRTYRGRLAVLPGRGSKTPASPVVVOGPDVAHLVP 240
 Db 181 SEKYRRLEIRFTVGTFTFRLASLRVQGLAVLPVGVASRRPAS-TLVQGPVDLHLVP 239

QY 241 LEEPYPVSHMTVPDEDFVLVLAHLHSHLGSEMFAPMGRCAGVNHLEFVYRAGVSRAML 300
 Db 240 LEEPYPVSHMTVPDEDFVLVLAHLHSHLGSEMFAPMGRCAGVNHLEFVYRAGVSRALL 299

QY 301 RLFLAMEGRHMEYECPLYVYPVYAFRLPEKDGKGVRAVDGELMAYSEAVQGVHPNFW 360
 Db 300 RLFLAMQKGMHLEDCPLYVYPVYAFRLPEKDGKGVRAVDGELMAYSEAVQGVHPNFW 359

QY 361 MVSGCVPPEPPPSKPOOMPPEEPL 383
 Db 360 MVCGSRDAPSGDRSGRPPPEEPL 382

RESULT 12
 ID AAY56054 standard; Protein: 388 AA.
 AC AAY56054;
 DT 28-MAR-2000 (first entry)
 DE Mouse sphingosine kinase 1b protein.
 KW Anticancer; antiproliferative; antiatherosclerotic; stroke; restenosis; antineurodegeneration; sphingosine kinase; cell proliferation; apoptosis; cancer; diabetic neuropathy; Alzheimer's disease; atherosclerosis.
 OS Mus musculus.
 PN WO9961581-A2.
 PD 02-DEC-1999.
 PF 25-MAY-1999; 99WO-US11521.
 PR 26-MAY-1998; 98US-0086657.
 PR 11-AUG-1998; 98US-0096049.
 PA (DEAN-) OFFICE DEAN RES & GRADUATE EDUCATION.

AAE07882
ID AAE07882 standard; Protein: 384 AA.
AC AAE07882;
XX
XX
XX 01-NOV-2001 (first entry)
XX
XX
XX Human sphingosine kinase (SphK) protein #1.
XX
XX Human; sphingosine kinase; SphK; restenosis; ischaemia; gene therapy;
XX anti-sense therapy; cancer; sphingolipid; signalling molecule; apoptosis;
XX cytostatic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;
XX leukemia; vasotropic; cell proliferative disorder; vascular disease.
XX
XX Homo sapiens.
XX
XX WO200160990-A2.
XX
XX 23-AUG-2001.
XX
XX 14-FEB-2001; 2001WO-US04789.
XX
XX 14-FEB-2000; 2000US-0182360.
XX 22-MAR-2000; 2000US-0191261.
XX
XX (CURA-) CURAGEN CORP.
XX (GENTH) GENTECH INC.
XX
XX Rastelli L;
XX
XX WPI: 2001-514770/56.
XX N-PSDB: AAD14424.
XX
XX An isolated sphingosine kinase polypeptide useful for treating a
XX PT SphK-associated disorder especially cancer, restenosis or ischemia in a
XX PT human -
XX
XX Claim 1: Page 90-91; 107pp; English.
XX
XX The present invention relates to sphingosine kinase (SphK) polypeptides
XX CC and nucleic acids encoding them. SphK is useful for treating a SphK-
XX CC associated disorder especially cancers such as leukemia, lymphoma,
XX CC ovarian, breast, lung colon, testicular, stomach and skin.
XX CC atherosclerosis, restenosis or ischemia and cell proliferative disease
XX CC or disorder associated with vascular diseases. SphK gene is used in gene
XX CC therapy and anti-sense therapy. Sphingolipids serving as signalling
XX CC molecules, have recently emerged as regulators of cell growth.
XX CC differentiation, diverse cell phenotypes and cell death. Activation of
XX CC SphK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
XX CC endothelial cells. The present sequence is human sphingosine kinase
XX CC (SphK) protein.
XX
XX Sequence 384 AA;
XX
XX Query Match 97.3%; Score 1962; DB 22; Length 384;
XX Best Local Similarity 97.7%; Pred. No. 1.1e-203;
XX Matches 375; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 241 LEEPVSHWTVPPEDFVLYLALLSHLGESEFAAPKRCACAGVHLFFYRAVSRAMLL 300
DB 241 LEEQVPSHWQVPPDEDFVLYLALLSHLASEMFAAPKRCACAGVHLFFYRAVSRAMLL 300
QY 301 RLFLAMERGRHMEYECPLYVVPVAFRLPEKDGKGFAYVDGELMVSEAVOGVHPNPFM 360
DB 301 RLFLAMERGRHMEYECPLYVVPVAFRLPEKDGKGFAYVDGELMVSEAVOGVHPNPFM 360
QY 361 MVSGCVPEPPSMKPOQMPPEEPL 384
DB 361 MVSGCVPEPPSMKPOQMPPEEPL 384
RESULT 10
ID ABB90209
XX ABB90209 standard; Protein: 333 AA.
XX
XX ABB90209;
XX
XX 24-MAY-2002 (first entry)
XX
XX Human polypeptide SEQ ID NO 2585.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX anti-allergic; hepatotoxic; antidiabetic; anti-inflammatory; anticancer;
XX vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.
XX
XX Homo sapiens.
XX
XX WO200190304-A2.
XX
XX 29-NOV-2001.
XX
XX 18-MAY-2001; 2001WO-US16450.
XX
XX 19-MAY-2000; 2000US-205515P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
XX
XX WPI: 2002-122018/16.
XX N-PSDB: ABL90618.
XX
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
XX PT prevention of neural, immune system, muscular, reproductive,
XX PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
XX PT disorders -
XX
XX Claim 11: SEQ ID NO 2585; 2081pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
XX CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
XX CC medical conditions e.g. by protein or gene therapy. The genes are
XX CC isolated from a range of human tissues disclosed in the specification.
XX CC The nucleic acids, proteins, antibodies and antagonists are useful
XX CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
XX CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX CC and parasitic infections.
XX CC Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at http://wipo.int/pub/published_pcl_sequences.

DR WPI: 2002-47846/51.
XX
PT New isolated sphingosine kinase, useful in identifying modulators for
PT treating e.g. cancer, also related nucleic acid, vectors and
PT transformed cells
XX
PS Disclosure: Fig 3; 24pp; English.
XX
CC The invention relates to an isolated sphingosine kinase (SPHK) DNA. Cells
CC transformed with SPHK DNA are used to screen for agents that reduce,
CC eliminate or promote SPHK activity. Agents that inhibit activity are
CC useful for decreasing cell proliferation, e.g. for treating cancer, and
CC for treating diseases associated with abnormal migration and motility of
CC cells, e.g. restenosis or diabetic neuropathy. Agents that increase
CC activity are used to reduce cell death. Antibodies raised against SPHK,
CC and primers or oligonucleotides derived from the DNA are useful for
CC diagnosis. The antibodies are also useful as therapeutic inhibitors. The
CC present sequence represents a human sphingosine kinase 1 (hSPHK1).
XX
SQ Sequence 384 AA:
XX
Query Match 99.6%; Score 2008; DB 23; Length 384;
Best Local Similarity 99.2%; Pred. No. 1,1e-208;
Matches 381; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 MDPAGGPRGVLPKRCRVLYLNPGRGKGAQLFRSHVQPLLAEEAISFTLMTERRNHA 60
DB 1 MDPAGGPRGVLPKRCRVLYLNPGRGKGAQLFRSHVQPLLAEEAISFTLMTERRNHA 60
QY 61 RELVASEELGRMDALVYVNSGDLMEHVNGLMERPDWETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVASEELGRMDALVYVNSGDLMEHVNGLMERPDWETAIOKPLCSLPAGSGNALAASL 120
QY 121 NHYAGEVQVNTEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSLAWGFADVDLE 180
DB 121 NHYAGEVQVNTEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSLAWGFADVDLE 180
QY 181 SEKVRRLGEMRTLTGTLRLAALRTYRGLAYLPVGRVSGKTPASPVVVOGQPVDAHLVP 240
DB 181 SEKVRRLGEMRTLTGTLRLAALRTYRGLAYLPVGRVSGKTPASPVVVOGQPVDAHLVP 240
QY 241 LEEPVPSSHTVPPDEDFVYLALLSHLSGSEMFAPMGRCACAGVNHLEFVVRAGVSRAAML 300
DB 241 LEEPVPSSHTVPPDEDFVYLALLSHLSGSEMFAPMGRCACAGVNHLEFVVRAGVSRAAML 300
QY 301 RLFLAMEKGRHMEYECPTLYVYVVAFRLEPKDGKGFVAVDGLMVSSEAVQGVHPNFTFW 360
DB 301 RLFLAMEKGRHMEYECPTLYVYVVAFRLEPKDGKGFVAVDGLMVSSEAVQGVHPNFTFW 360
QY 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
DB 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
XX
RESULT 8
AA96057
ID AA96057 standard; Protein; 384 AA.
XX
AC AA96057;
XX
DT 05-DEC-2000 (first entry)
XX
DE Human sphingosine kinase A.
XX
KW Sphingosine kinase A; SKA; human; drug screening; infection;
KW antiinflammatory; antiallergic; anticancer; inflammation; allergy;
KW cancer; therapy; diagnosis.
XX
OS Homo sapiens.
XX
PN MO200052173-A2.
XX
PD 08-SEP-2000.

XX
PF 02-MAR-2000; 2000W0-CA00223.
XX
PR 02-MAR-1999; 99US-0122516.
XX
PA (ALLX) NPS ALLELIX CORP.
XX
PI Munroe D, Gupta A, Falzone GR;
XX
DR WPI: 2000-572185/53.
DR N-PSDB; AAA50508.
XX
PT New human sphingosine kinase A, B and C polynucleotides and
PT polypeptides useful in e.g. chromosome and gene mapping, and detecting
PT inflammation or disease associated with abnormal levels of sphingosine
PT kinase expression
XX
PS Disclosure: Fig 3; 81pp; English.
XX
XX The present sequence is that of human sphingosine kinase A (SKA),
CC an enzyme that phosphorylates sphingosine to form sphingosine
CC 1-phosphate. The sequence was deduced from that of a polynucleotide
CC (see AAA50508) isolated from an HeLa cDNA library. The invention
CC provides polynucleotides (see AAA50508-10) and polypeptides (see
CC AA96057-59) for the human sphingosine kinase (SK) homologues SKA,
CC SKB and SKC. The polypeptides can be obtained using recombinant
CC DNA methods, and host cells containing expression vectors including
CC SK polynucleotides are used in a claimed method of screening for
CC compounds that inhibit or activate human SK activity. Human SK
CC specific antibodies, inhibitors, ligands or their analogues can be
CC used as bioactive agents to treat inflammation or disease including
CC viral, bacterial or fungal infections, allergic responses,
CC mechanical injury associated with trauma, hereditary diseases,
CC lymphoma or carcinoma, and other conditions with activate the
CC genes of kidney, lung, heart, lymphoid or tissues of the nervous
CC system.
XX
SQ Sequence 384 AA:
XX
Query Match 99.2%; Score 2000; DB 21; Length 384;
Best Local Similarity 99.2%; Pred. No. 8e-208;
Matches 381; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1 MDPAGGPRGVLPKRCRVLYLNPGRGKGAQLFRSHVQPLLAEEAISFTLMTERRNHA 60
DB 1 MDPAGGPRGVLPKRCRVLYLNPGRGKGAQLFRSHVQPLLAEEAISFTLMTERRNHA 60
QY 61 RELVASEELGRMDALVYVNSGDLMEHVNGLMERPDWETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVASEELGRMDALVYVNSGDLMEHVNGLMERPDWETAIOKPLCSLPAGSGNALAASL 120
QY 121 NHYAGEVQVNTEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSLAWGFADVDLE 180
DB 121 NHYAGEVQVNTEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSLAWGFADVDLE 180
QY 181 SEKVRRLGEMRTLTGTLRLAALRTYRGLAYLPVGRVSGKTPASPVVVOGQPVDAHLVP 240
DB 181 SEKVRRLGEMRTLTGTLRLAALRTYRGLAYLPVGRVSGKTPASPVVVOGQPVDAHLVP 240
QY 241 LEEPVPSSHTVPPDEDFVYLALLSHLSGSEMFAPMGRCACAGVNHLEFVVRAGVSRAAML 300
DB 241 LEEPVPSSHTVPPDEDFVYLALLSHLSGSEMFAPMGRCACAGVNHLEFVVRAGVSRAAML 300
QY 301 RLFLAMEKGRHMEYECPTLYVYVVAFRLEPKDGKGFVAVDGLMVSSEAVQGVHPNFTFW 360
DB 301 RLFLAMEKGRHMEYECPTLYVYVVAFRLEPKDGKGFVAVDGLMVSSEAVQGVHPNFTFW 360
QY 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
DB 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
XX
RESULT 9

```

Db      121 NHYAGYEQVNTEDLLTNCITLLCRRLISPMNLLSLHTASGLRFSVLSLAWGFIADVLE 180
QY      181 SEKYRRLSEMFTLTGTEFLRLAALRTYRGRLAYLPVGRVGSKTTPASPVVVOGCPVDALHP 240
Db      181 SEKYRRLSEMFTLTGTEFLRLAALRTYRGRLAYLPVGRVGSKTTPASPVVVOGCPVDALHP 240
QY      241 LEEPVPSHMTVVPDEDFVLVALLSHLCSSEMFPAAPMGRCAAGVWHLFYVRAGVSRAMLL 300
Db      241 LEEPVPSHMTVVPDEDFVLVALLSHLCSSEMFPAAPMGRCAAGVWHLFYVRAGVSRAMLL 300
QY      301 RLFLAMEGRHMEYECPLYVVPVAFRLPEPKDGKGVFAVNDGELMVSEAVOGVHPNYFW 360
Db      301 RLFLAMEGRHMEYECPLYVVPVAFRLPEPKDGKGVFAVNDGELMVSEAVOGVHPNYFW 360
QY      361 WVSGCVPPEPPSMKPOQMPPEEPL 384
Db      361 WVSGCVPPEPPSMKPOQMPPEEPL 384

RESULT 6
AA040180
ID      AA040180 standard; protein; 384 AA.
XX
AC      AA040180;
XX
XX      22-OCT-2001 (first entry)
XX
DE      Human polypeptide SEQ ID NO 3325.
XX
KW      Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW      peripheral nervous system; neuropathy; central nervous system; CNS;
KW      Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW      amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW      chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW      leukaemia.
XX
OS      Homo sapiens.
XX
PN      WO200153312-A1.
XX
PD      26-JUL-2001.
XX
PF      26-DEC-2000; 2000MO-US34263.
XX
PR      21-JAN-2000; 2000US-0488725.
PR      25-APR-2000; 2000US-052317.
PR      09-JUL-2000; 2000US-0598042.
PR      19-JUL-2000; 2000US-0620312.
PR      03-AUG-2000; 2000US-0653450.
PR      14-SEP-2000; 2000US-0662191.
PR      19-OCT-2000; 2000US-0693036.
PR      29-NOV-2000; 2000US-0727344.
XX
PA      (HYSE-) HYSEQ INC.
XX
PI      Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI      Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI      Zhao QA, Zhou P, Goodrich R, Dirmacac RT;
XX
DR      WPI; 2001-442253/47.
DR      N-PSDB; AAI59336.
XX
PT      Novel nucleic acids and polypeptides, useful for treating disorders
XX      such as central nervous system injuries -
XX
XX      Example 5: SEQ ID NO 3325; 10078pp; English.
XX
XX      The invention relates to human nucleic acids (AA157798-AA161369) and
XX      the encoded polypeptides (AA038642-AA042213) with nocotropic,
XX      immunosuppressant and cytostatic activity. The polynucleotides are useful
XX      in gene therapy. A composition containing a polypeptide or polynucleotide
XX      of the invention may be used to treat diseases of the peripheral nervous
XX      system, such as peripheral nervous injuries, peripheral neuropathy and

```

```

CC      localized neuropathies and central nervous system diseases, such as
CC      Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC      lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC      utilisation of the activities such as: immune system suppression,
CC      activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC      and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC      assays for receptor activity, arthritis and inflammation, leukaemias and
CC      C.N.S disorders.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification.
XX
SQ      Sequence 384 AA:
XX
Query Match          99.8%; Score 2011; Db 22; Length 384;
Best Local Similarity 99.7%; Pred. No. 5.1e-209;
Matches 383; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY      1 MDPACGPRGVLPKRCVRLVILNPRGCKAKQLQFRSHVQPLAEAEISFTLMTERRNA 60
Db      1 MDPACGPRGVLPKRCVRLVILNPRGCKAKQLQFRSHVQPLAEAEISFTLMTERRNA 60
QY      61 RELVASEELGRMDALVVMSSDGLMHEVNGLMERPDWETAIOKPLCSLPGSGNALAPSL 120
Db      61 RELVASEELGRMDALVVMSSDGLMHEVNGLMERPDWETAIOKPLCSLPGSGNALAPSL 120
QY      121 NHYAGYEQVNTEDLLTNCITLLCRRLISPMNLLSLHTASGLRFSVLSLAWGFIADVLE 180
Db      121 NHYAGYEQVNTEDLLTNCITLLCRRLISPMNLLSLHTASGLRFSVLSLAWGFIADVLE 180
QY      181 SEKYRRLSEMFTLTGTEFLRLAALRTYRGRLAYLPVGRVGSKTTPASPVVVOGCPVDALHP 240
Db      181 SEKYRRLSEMFTLTGTEFLRLAALRTYRGRLAYLPVGRVGSKTTPASPVVVOGCPVDALHP 240
QY      241 LEEPVPSHMTVVPDEDFVLVALLSHLCSSEMFPAAPMGRCAAGVWHLFYVRAGVSRAMLL 300
Db      241 LEEPVPSHMTVVPDEDFVLVALLSHLCSSEMFPAAPMGRCAAGVWHLFYVRAGVSRAMLL 300
QY      301 RLFLAMEGRHMEYECPLYVVPVAFRLPEPKDGKGVFAVNDGELMVSEAVOGVHPNYFW 360
Db      301 RLFLAMEGRHMEYECPLYVVPVAFRLPEPKDGKGVFAVNDGELMVSEAVOGVHPNYFW 360
QY      361 WVSGCVPPEPPSMKPOQMPPEEPL 384
Db      361 WVSGCVPPEPPSMKPOQMPPEEPL 384

RESULT 7
AB08089
ID      AB08089 standard; protein; 384 AA.
XX
AC      AB08089;
XX
XX      10-SEP-2002 (first entry)
XX
DE      Human sphingosine kinase 1 (hSPHK1) protein sequence.
XX
KW      Sphingosine kinase; SPHK; SPHK1; cytostatic; vasotropic; antidiabetic;
KW      neuroprotective; human; enzyme.
XX
OS      Homo sapiens.
XX
PN      US2002042358-A1.
XX
PD      11-APR-2002.
XX
PF      02-MAR-2001; 2001US-0796487.
XX
PR      02-MAR-2000; 2000US-186352P.
XX
PA      (SPIE/) SPIEGEL S.
XX
PI      Spiegel S;
XX

```

KW T helper-1 related disease; chronic obstructive pulmonary disease;
 KW asthma; myocardial infarction; neurodegenerative disorder;
 KW wound healing; embryogenesis; anticoagulant; cerebroprotective;
 KW neuroprotective; antipruritic; antiarthritic; cytostatic; cardiant;
 KW vulnerable.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT Region 16..95
 FT /label= Conserved_region
 PN WC200131029-A2.
 PD
 XX 03-MAY-2001.
 XX
 PF 27-OCT-2000; 2000MO-EP09498.
 XX
 PR 28-OCT-1999; 99US-0162307.
 PR 07-FEB-2000; 2000US-0180525.
 XX
 PA (WARN) WARNER LAMBERT CO.
 PI Allen J, Goslink M, Melendez AJ, Takacs L;
 DR WPI: 2001-300510/31.
 DR N-PSDB; AAD04477.
 XX
 PT New human sphingosine kinase type I gene for screening drug candidates
 PT particularly inhibitors used for preventing or treating e.g.
 PT atherosclerosis, thrombosis, asthma and diabetes
 XX
 PS Claim 3; Fig 1; 91pp; English.
 XX
 CC The present sequence is human sphingosine kinase type I (hsk1).
 CC The hsk1 gene is located on chromosome 17q25.2. The sk1 converts the
 CC substrate sphingosine to sphingosine-1-phosphate (S1P). The sk1 gene
 CC and encoded polypeptide are applicable in screening drug candidates
 CC particularly inhibitors for preventing or treating disorders such as
 CC haemostasis, thrombosis, allergic reactions, proliferative diseases
 CC including cancer, haematopoietic disorders such as leukaemia,
 CC cardiovascular diseases such as stroke, atherosclerosis and coronary
 CC artery disease, dyslipidaemia, diabetes including type I and type II
 CC diabetes, autoimmune and inflammatory diseases such as multiple
 CC sclerosis, T helper-1 related diseases, chronic obstructive pulmonary
 CC disease, asthma, myocardial infarction, neurodegenerative disorders,
 CC natural wound healing processes and embryogenesis.
 CC
 XX
 SO Sequence 384 AA;
 Query Match 100.0%; Score 2016; DB 22; Length 384;
 Best Local Similarity 100.0%; Pred. No. 1.5e-209;
 Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDPAGGPRGVLPDRCRVYLLNPRGKGKALQFRSHVQPLLAEAETSTLMTERRNHA 60
 DB 1 MDPAGGPRGVLPDRCRVYLLNPRGKGKALQFRSHVQPLLAEAETSTLMTERRNHA 60
 QY 61 RELVSEELGKMDALVWNSGDLMEHVNGLMERPDWETAIOKPLCSIPASSGNALASL 120
 DB 61 RELVSEELGKMDALVWNSGDLMEHVNGLMERPDWETAIOKPLCSIPASSGNALASL 120
 QY 121 NHVAGYEDVNTEDLLTNCITLLCRRLSPMNLISLHTASGLRFSVSLANGFTADVLE 180
 DB 121 NHVAGYEDVNTEDLLTNCITLLCRRLSPMNLISLHTASGLRFSVSLANGFTADVLE 180
 QY 181 SEKYRRLGEMFTLTGTLRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
 DB 181 SEKYRRLGEMFTLTGTLRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
 QY 241 LEEPVSHWTVVPPDDFVLYVALLSHLSGEMFAAPMGRCAGVNHLLFYVAGVSRAML 300
 DB 241 LEEPVSHWTVVPPDDFVLYVALLSHLSGEMFAAPMGRCAGVNHLLFYVAGVSRAML 300

QY 301 RLFLAMEKGRHMEYECYLYVVPVAFRLPEPKDGYFAVDGELMSEAVOGVHPVFW 360
 DB 301 RLFLAMEKGRHMEYECYLYVVPVAFRLPEPKDGYFAVDGELMSEAVOGVHPVFW 360
 QY 361 MVSGVEPPPSWKPQOMPPEEPL 384
 DB 361 MVSGVEPPPSWKPQOMPPEEPL 384
 RESULT 5
 AAB48007
 ID AAB48007 standard; Protein: 384 AA.
 XX
 AC AAB48007;
 XX
 DT 19-MAR-2001 (first entry)
 XX
 DE Human sphingosine kinase protein sequence.
 XX
 KW Sphingosine kinase; antiarthritic; antiasthmatic; antiarteriosclerotic;
 KW antiinflammatory; neuroprotective; antibacterial; immunosuppressive;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN WO200070028-A1.
 XX
 PD 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000MO-AU00457.
 XX
 PR 13-MAY-1999; 99AU-0000339.
 PR 08-JUL-1999; 99AU-0001504.
 XX
 PA (JOHU) JOHNSON & JOHNSON RES PTY LTD.
 PI Pltson SM, Wattenberg BW, Xia P, D'Andrea RJ, Gamble JR, Vades MA;
 DR WPI: 2001-016227/02.
 DR N-PSDB; AAC84161.
 XX
 PT Novel sphingosine kinase protein and nucleic acid molecules for
 PT diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma,
 PT atherosclerosis, inflammation, meningitis, multiple sclerosis and
 PT septic shock
 XX
 PS Claim 9; Fig 7a; 100pp; English.
 XX
 CC This represents a human sphingosine kinase (SK) protein. The human SK
 CC protein, encoding nucleic acids and modulators are useful for modulating
 CC expression, functional activity or cellular functional activity of
 CC sphingosine kinase in a subject and also for treating a mammal by
 CC modulating the activity of SK. Diseases treated by regulating SK
 CC cellular activity include rheumatoid arthritis, asthma, atherosclerosis,
 CC inflammation, meningitis, multiple sclerosis and septic shock.
 CC
 XX
 SO Sequence 384 AA;
 Query Match 99.9%; Score 2013; DB 22; Length 384;
 Best Local Similarity 99.7%; Pred. No. 3.1e-209;
 Matches 383; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDPAGGPRGVLPDRCRVYLLNPRGKGKALQFRSHVQPLLAEAETSTLMTERRNHA 60
 DB 1 MDPAGGPRGVLPDRCRVYLLNPRGKGKALQFRSHVQPLLAEAETSTLMTERRNHA 60
 QY 61 RELVSEELGKMDALVWNSGDLMEHVNGLMERPDWETAIOKPLCSIPASSGNALASL 120
 DB 61 RELVSEELGKMDALVWNSGDLMEHVNGLMERPDWETAIOKPLCSIPASSGNALASL 120
 QY 121 NHVAGYEDVNTEDLLTNCITLLCRRLSPMNLISLHTASGLRFSVSLANGFTADVLE 180
 DB 121 NHVAGYEDVNTEDLLTNCITLLCRRLSPMNLISLHTASGLRFSVSLANGFTADVLE 180

Best Local Similarity 100.0%; Pred. No. 1.5e-209;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGPGCVLPKRCVLYLNLNPGGCKALQLFKSHVOPLLAEAFISFTLMLTERRNA 60
Db 1 MDPAGGPGCVLPKRCVLYLNLNPGGCKALQLFKSHVOPLLAEAFISFTLMLTERRNA 60
QY 61 RELVSEELGEMFRLGTEFLRLAALRTYRGLAYLPVGRVSKTPASPVVQGGVDAHLVP 240
Db 61 RELVSEELGEMFRLGTEFLRLAALRTYRGLAYLPVGRVSKTPASPVVQGGVDAHLVP 240
QY 121 NHYAGYEOVTNEDLLTNCITLLCRLLSPPMNLISLHTASGLRFLSVLSLAWGFIADVLE 180
Db 121 NHYAGYEOVTNEDLLTNCITLLCRLLSPPMNLISLHTASGLRFLSVLSLAWGFIADVLE 180
QY 181 SEKRYRLGEMFRLGTEFLRLAALRTYRGLAYLPVGRVSKTPASPVVQGGVDAHLVP 240
Db 181 SEKRYRLGEMFRLGTEFLRLAALRTYRGLAYLPVGRVSKTPASPVVQGGVDAHLVP 240
QY 241 LEEPVPSHMTVVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVMHLFVYRAGVSRAML 300
Db 241 LEEPVPSHMTVVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVMHLFVYRAGVSRAML 300
QY 301 RLFLAMEKGRHMEYECFYLVPVVAFLRLEPKDGKGFVAVDGLMSEAVOQGVHNFWM 360
Db 301 RLFLAMEKGRHMEYECFYLVPVVAFLRLEPKDGKGFVAVDGLMSEAVOQGVHNFWM 360
QY 361 MWSGCVPEPPSMKPOQMPPEEPL 384
Db 361 MWSGCVPEPPSMKPOQMPPEEPL 384

RESULT 3
AAB94589 standard; Protein: 384 AA.

QY 26-JUN-2001 (first entry)
Db 26-JUN-2001 (first entry)

Human protein sequence SEQ ID NO:15394.
Human: primer: detection; diagnosis; antisense therapy; gene therapy.
Homo sapiens.
EPI074617-A2.
07-FEB-2001.
28-JUL-2000; 2000EP-0116126.
29-JUL-1999; 99JP-0248036.
27-AUG-1999; 99JP-0300253.
11-JAN-2000; 2000JP-0118776.
02-MAY-2000; 2000JP-0183767.
09-JUN-2000; 2000JP-0241899.
(HELI-) HELIX RES INST.
Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
WPI: 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -
Claim 8: SEQ ID 15394; 2537pp + CD ROM, English.
The present invention describes primer sets for synthesizing 5602

full-length cDNAs defined in the specification, where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to a 5'-end complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers used in the method of the full-length cDNAs easily without any specialised method. AAH03166 to AAH13628 and AAH3633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH93693 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 384 AA:
SQ
Query Match 100.0%; Score 2016; DB 22; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.5e-209;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGPGCVLPKRCVLYLNLNPGGCKALQLFKSHVOPLLAEAFISFTLMLTERRNA 60
Db 1 MDPAGGPGCVLPKRCVLYLNLNPGGCKALQLFKSHVOPLLAEAFISFTLMLTERRNA 60
QY 61 RELVSEELGEMFRLGTEFLRLAALRTYRGLAYLPVGRVSKTPASPVVQGGVDAHLVP 240
Db 61 RELVSEELGEMFRLGTEFLRLAALRTYRGLAYLPVGRVSKTPASPVVQGGVDAHLVP 240
QY 121 NHYAGYEOVTNEDLLTNCITLLCRLLSPPMNLISLHTASGLRFLSVLSLAWGFIADVLE 180
Db 121 NHYAGYEOVTNEDLLTNCITLLCRLLSPPMNLISLHTASGLRFLSVLSLAWGFIADVLE 180
QY 181 SEKRYRLGEMFRLGTEFLRLAALRTYRGLAYLPVGRVSKTPASPVVQGGVDAHLVP 240
Db 181 SEKRYRLGEMFRLGTEFLRLAALRTYRGLAYLPVGRVSKTPASPVVQGGVDAHLVP 240
QY 241 LEEPVPSHMTVVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVMHLFVYRAGVSRAML 300
Db 241 LEEPVPSHMTVVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVMHLFVYRAGVSRAML 300
QY 301 RLFLAMEKGRHMEYECFYLVPVVAFLRLEPKDGKGFVAVDGLMSEAVOQGVHNFWM 360
Db 301 RLFLAMEKGRHMEYECFYLVPVVAFLRLEPKDGKGFVAVDGLMSEAVOQGVHNFWM 360
QY 361 MWSGCVPEPPSMKPOQMPPEEPL 384
Db 361 MWSGCVPEPPSMKPOQMPPEEPL 384

RESULT 4
AAB00924 standard; Protein: 384 AA.

QY 04-JUL-2001 (first entry)
Db 04-JUL-2001 (first entry)

Human sphingosine kinase type 1 (hsk1).
Human: sphingosine kinase type 1; sk1; chromosome 17q25.2;
sphingosine-1-phosphate; S1P; drug screening; therapy; haemostasis;
thrombosis; allergic reaction; proliferative disease; cancer;
haematopoietic disorder; leukaemia; cardiovascular disease; stroke;
atherosclerosis; coronary artery disease; dyslipidaemia; diabetes;
autoimmune disease; inflammatory disease; multiple sclerosis;

QY 345 MYSEAV 350
 Db 334 VLKRAV 339

RESULT 11
 SM4B_MOUSE
 ID SM4B_MOUSE STANDARD: PRT: 782 AA.
 AC 062179:
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Semaphorin 4B (Semaphorin C) (Sema C) (Fragment).
 GN SEMA4B OR SEMAC OR SEMC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRI; TISSUE=Brain;
 RX MEDLINE=95267431; PubMed=7748561;
 RA Puschel A.W., Adams R.H., Betz H.;
 RT "Murine semaphorin D/collapsin is a member of a diverse gene family
 RT and creates domains inhibitory for axonal extension.";
 RL Neuron 14:941-948(1995).
 RN [2]
 RP INTERACTION WITH GIPC.
 RX MEDLINE=99253973; PubMed=10318831;
 RA Wang L.-H., Kalb R.G., Strittmatter S.M.;
 RT "A PDZ protein regulates the distribution of the transmembrane
 RT semaphorin, M-SemF.";
 RL J. Biol. Chem. 274:14137-14146(1999)
 CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
 CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
 CC -1- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
 CC LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH
 CC MODERATE LEVELS FROM THEN UNTIL BIRTH.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X85992; CAA59984.1; .
 CC MGD: MGI:107559; Sema4b.
 CC InterPro: IPR003659; Plexin-like.
 CC InterPro: IPR002165; Plexin_repeat.
 CC InterPro: IPR001627; Sema.
 CC Pfam: PF01403; Sema; 1.
 CC SMART: SM00423; PSI; 1.
 CC Transmembrane: Immunoglobulin domain; Multigene family; Neurogenesis;
 CC Developmental protein; Glycoprotein.
 CC NON_TER 1
 CC DOMAIN 1
 CC TRANSMEM 663 662
 CC DOMAIN 663 683
 CC DOMAIN 684 782
 CC DOMAIN 202 492
 CC DOMAIN 548 608
 CC DOMAIN 703 726
 CC DIOLFID 553 601
 CC CAROHYD 12 12
 CC CAROHYD 15 15

FT CAROHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 526 526 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 574 574 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHYD 639 639 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 782 AA; 86823 MW; 627A81FC8B87AC8 CRC64;

Query Match 4.6%; Score 92; DB 1; Length 782;
 Best Local Similarity 19.1%; Pred. No. 41; Indels 138; Gaps 19;
 Matches 81; Conservative 49; Mismatches 156;

QY 29 KALQFRSHVQPLAEAEISFTLTERNRHARELYRSEELG---RMDALVYMSGDGLMH 85
 Db 364 RVLNFLKDH---FLMDGVRSKLLLDQPARQYRAVHVRPLGHSYVDVLFCTGGRH 420
 QY 86 EVVNGIMERPDMETAIO-----KPLCSLPAGSGMALASLNHYAGEOYTNEDLLTCTL 140
 Db 421 KAVT-LSSRVHIIIELOIFPOGOPYONLLDSHGILYASSH-SGVQVP---VANCSL 474
 QY 141 L-LCRRLSPMNLISLHTSGLRFS-----VLSLWGFADYDSEKRYRIGEMRTL 194
 Db 475 YPTGDDLARPYCAMTGSACRLASLYOPDLASRPW--TDEGASVK----- 521
 QY 195 GTFELRALTYRGRLAYLPVGRVSKTPASPVYVOGPDVLAHVLEPPVPSH-WT--- 250
 Db 522 ----ELCKNSSTKAR--FLVPGK-----FCKQYQIOPNTYNTIACLLSLATRLVHNG 570
 QY 251 -----VPEDEDFVLVALLSHLSGSEMFAPMGCAACVMHLFVYRACVSAMLL 300
 Db 571 APVNASAGRVLPPTGDLILV-----GSOQ-----GLGVFCWSTIEEGFOOLVAS 614
 QY 301 RLFLAMEKG-----RHNIECEPYLYVP 323
 Db 615 YCEYVEEGMDQKNOQDGTPIIINTSRVASAPAGRDSMGADKSYNNEELVMTLFEVFM 674
 QY 324 VVAFRF---EPKDKGVFAVDCELAMVSEAVOGVHNFMVMSGVEPPSPMKPP 380
 Db 675 VLLFLEFLYRHRGMLFLKQGE-----CASVHKTRPYLVLP 713

QY 381 EEPL 384
 Db 714 TRPL 717

RESULT 12
 NAL1_HUMAN
 ID NAL1_HUMAN STANDARD: PRT: 1473 AA.
 AC 09C000: Q9Y2E0; Q9HAV8; Q9BZ29; Q9BZ28; Q9UFT4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE NACHT-, LRR- and PYD-containing protein 2 (Death effector filament-
 DE forming ced-4-like apoptosis protein) (Nucleotide-binding domain and
 DE caspase recruitment domain) (Caspase recruitment domain protein 7).
 GN NALP1 OR DEFCAP OR NAC OR CARD7 OR KIAA0926.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21169419; PubMed=11270363;
 RA Bertin J., Distefano P.S.;
 RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
 RT proteins.";
 RN Cell Death Differ. 7:1273-1274(2000).
 RP [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21148093; PubMed=11250163;
 RA Martiny-Baron F., Hofmann K., Tschopp J.;
 RT "The pyrin domain: a possible member of the death domain-fold family

RT Implicated in apoptosis and inflammation.";
 RL Curr. Biol. 11:R118-R120(2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE-Erythroleukemia.
 RX MEDLINE-21153743; PubMed-11076957;
 RA Hlaing T., Guo R.-F., Dilley K.A., Loussia J.M., Morrish T.A.,
 RA Shi M.M., Vincenz C., Ward P.A.;
 RT Molecular cloning and characterization of DEFCAP-L and -S, two
 RT isoforms of a novel member of the mammalian Ced-4 family of apoptosis
 RT proteins.";
 RL J. Biol. Chem. 276:9230-9238(2001).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND PROTEIN INTERACTION.
 RC TISSUE-T-cell;
 RX MEDLINE-21153744; PubMed-1113115;
 RA Chu Z.-L., Pio F., Xie Z., Welsh K., Krajewski M., Krajewski S.,
 RA Godzik A., Reed J.C.;
 RT "A novel enhancer of the Apaf1 apoptosome involved in cytochrome
 RT c-dependent caspase activation and apoptosis.";
 RL J. Biol. Chem. 276:9239-9245(2001).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-Brain;
 RX MEDLINE-99246063; PubMed-10231032;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohata O.;
 RT Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:63-70(1999).
 RN [6]
 RP SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).
 RC TISSUE-Uterus;
 RA Koehler K., Beyer A., Mewes H.-W., Gassenhuber J., Wleemann S.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: Able to form cytoplasmic structures termed death
 CC effector filaments. Enhances Apaf1 and cytochrome c-dependent
 CC activation of pro-caspase-9 and consecutive apoptosis.
 CC Seems to bind ATP.
 CC - SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9
 CC and with Apaf1 in a cytochrome c-inducible way leading to the
 CC formation of an apoptosome. This interaction may be ATP-dependent.
 CC - SUBCELLULAR LOCATION: Cytoplasmic.
 CC - ALTERNATIVE PRODUCTS: 4 isoforms: 1/NAC beta/DEFCAP-L (shown
 CC here), 2/NAC alpha/DEFCAP-S, 3/NAC gamma and 4/NAC delta; are
 CC produced by alternative splicing.
 CC - TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are
 CC expressed in peripheral blood leukocytes, chronic myelogenous
 CC leukemia cell line K-562, followed by thymus, spleen and heart.
 CC Also detected in lung, placenta, small intestine, colon, kidney,
 CC liver and muscle.
 CC - SIMILARITY: CONTAINS 1 DAPIN DOMAIN.
 CC - SIMILARITY: CONTAINS 1 NACHT DOMAIN.
 CC - SIMILARITY: CONTAINS 1 CARD DOMAIN.
 CC - SIMILARITY: CONTAINS 7 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR HSP: P13489; 1AAY.
 DR MIM: 606636;
 DR InterPro: IPR00767; Disease_resist.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003590; LRR_RNinh.
 DR InterPro: IPR004020; PAAD_DAPIN_dom.
 DR Pfam: PF00560; LRR_1.
 DR Pfam: PF02758; PAAD_DAPIN_1.
 DR PRINTS: PR00364; DISEASERISIT.
 DR SMART: SM00368; LRR_R1_5.
 DR PROSITE: PS50209; CARD.
 DR PROSITE: PS50824; DAPIN_1.
 DR PROSITE: PS50837; NACHT.
 DR Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
 KW Alternative splicing.
 KM
 FT DOMAIN 1 92
 FT REPEAT 328 637
 FT REPEAT 704 725
 FT REPEAT 807 830
 FT REPEAT 864 887
 FT REPEAT 921 944
 FT REPEAT 950 973
 FT REPEAT 1199 1215
 FT REPEAT 1216 1236
 FT DOMAIN 1374 1463
 FT NP_BIND 334 341
 FT VARSPLIC 958 987
 FT VARSPLIC 1262 1305
 FT MUTAGEN 340 340
 FT CONFLICT 155 155
 FT CONFLICT 246 246
 FT CONFLICT 782 782
 FT CONFLICT 878 878
 FT CONFLICT 995 995
 FT CONFLICT 1119 1119
 FT CONFLICT 1184 1184
 FT CONFLICT 1241 1241
 FT CONFLICT 1366 1366
 SQ SEQUENCE 1473 AA: 165865 MW: 438FDDCE45C2562D CRC64;
 Query Match 4.68; Score 92; DB 1; Length 1473;
 Best local Similarity: 23.28; Pred No. 9.4; Mismatches 107; Indels 146; Gaps 21;
 Matches 89; Conservative 42;
 Db 78 MSGDGLMEHVVGIMERDMEFALOKPLSLP-----AGSG-----NLAASLHHVAGYE 127
 Db 815 LSGNSLSHSAVASKL-----TLRRPCLLETLRAGGGLTAEDCKDLAFGLR--ANO 865
 Oy 128 QVTNEDLTLNC-----TLILCRLLSP-----MNLISLHTASGL--RLSSVLSLAWGFI 174
 Db 866 TLTELDLSPNVLTGAGAKNHLQRLQPSCKLQRLQVSCGLTSDCCDLASVLS-ASPSL 924
 Oy 175 ADVVLESE-----KYRRLEGEMRFTLGTFLR--LAALRTVNGRL-- 210
 Db 925 KELDLOONNLDGVGRLLCEGLRHPACKLIRGLDQTLSDRMQELRALQEKPPOLLIF 984
 Oy 211 -----AYLVGRVGSKTPASPVVVOGQPVDAHLVPL----- 241
 Db 985 SRKPSVMTPTGGLDTGEMSNSTSLKRQRLGSGERASAVH-----AOANLKLDVSKIF 1038
 Oy 242 -----EEPVHMTVVVDEDFVYLAL-----LISH-LGS-EWFAAPMGRCAGVW 285
 Db 1039 PLAEIAEESPE--VVVE-----LILCVSPASQGLIHKPLGTDDDFWPTGVPVATEV- 1090
 Oy 286 HLFYVAGVSRMLRLRLFLAMEGRHNEYECPLVYVPAVFRLEPRDG-----KGFVAV 340
 Db 1091 -----VDKRNK-----YRVHFPVAGSYRNPNTGLCFVWREAVIV 1125
 Oy 341 DGLWAVEAVGOVHPNFWVSG 364
 Db 1126 EIEFCWDOFLGEINPOHSMWVAG 1149

RESULT 13
PR36_CAEEL STANDARD: PRT: 414 AA.
AC P46502:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Probable 26S protease regulatory subunit 6S.
GN F23F12.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Du Z.
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE 26S PROTEASE IS INVOLVED IN THE ATP-DEPENDENT
CC DEGRADATION OF UBQUITINATED PROTEINS. THE REGULATORY (OR ATPASE)
CC COMPLEX CONTERS ATP DEPENDENCY AND SUBSTRATE SPECIFICITY TO THE
CC 26S COMPLEX (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U12965; AAA20608.1; -
DR Wormpep; F23F12.6; CE01253.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003959; AAA_ATPase_centre.
DR InterPro: IPR003960; AAA_sub.
DR Pfam: PF00004; AAA.1.
DR SMART: SM00382; AAA.1.
DR TIGRFAMs: TIGR01242; 26Sp45; 1.
DR PROSITE: PS00674; AAA.1.
KW Hypothetical protein; Proteasome; ATP-binding; Nuclear protein.
FT NP_BIND 202 209 ATP (POTENTIAL).
SQ SEQUENCE 414 AA; 46358 MM; 8630AFB8A2C7FF32 CRC64;
Query Match 4.48; Score 89; DB 1; Length 414;
Best Local Similarity 22.08; Pred. No. 3.2;
Matches 46; Conservative 27; Mismatches 78; Indels 58; Gaps 5;
OY 42 LAEAFSTLTALTEERRNARELVSSEELGRMDALVVMGSDGLMEHYVGLMERPMETAI 101
DB 47 LAHLOVMEDYIKLETRNLKEELHNAQE-----EV----- 75
OY 102 QKPLCSLPAGSGNALAASLNHYAGYQVNTNEDLLTNCILLCRLLSPNNLLSLHTA 161
DB 76 -KRISQVPLVIGQFLAENDQNAIYGTSGSNYYRVLSILDRLEKRGCSVALHKYEN- 133
OY 162 RLFSVL-----SLMGFIADVDESSEKRYRLGEMRTTLGTFLELAALR 204
DB 134 ALVYLPLEADSSIQMLRPDEKPDLSYGDIGLDMQKOEVRAYELPLTHGELYOOIGID 193
OY 205 TYRGLIATLPVG-----RVSGKTSPAS 225
DB 194 PPRGVLMYGPCCGKTMIAKAVAAVTAAS 222
RESULT 14
SERA_RAT STANDARD: PRT: 533 AA.
ID SERA_RAT
AC 008651:
DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (3-PGDH).
GN PGDH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97250633; PubMed=9163325;
RA Achouri Y., Rider M.H., van Schaftingen E., Robbi M.;
RT "Cloning, sequencing and expression of rat liver 3-phosphoglycerate
RT dehydrogenase.";
RL Biochem J. 323:365-370(1997).
CC -1- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) = 3-
CC phosphohydroxypyruvate + NADH.
CC -1- PATHWAY: Serine biosynthesis; first step.
CC -1- SUBUNIT: Homotrimer.
CC -1- TISSUE SPECIFICITY: LIVER, KIDNEY, BRAIN, TESTIS.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X97772; CAA66374.1; -
DR HSSP: P08328; 1PSD.
DR InterPro: IPR002162; D_2hyd-ac-dh.
DR Pfam: PF00389; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR PROSITE: PS00065; D_2-HYDROXYACID_DH.1; 1.
DR PROSITE: PS00670; D_2-HYDROXYACID_DH.2; 1.
DR PROSITE: PS00671; D_2-HYDROXYACID_DH.3; 1.
KW Serine biosynthesis; Oxidoreductase; NAD.
FT ACT_SITE 236 236 SUBSTRATE-BINDING (BY SIMILARITY).
FT ACT_SITE 265 265 BY SIMILARITY.
FT ACT_SITE 283 283 BY SIMILARITY.
SQ SEQUENCE 533 AA; 56493 MM; 7273DAC3349595EP CRC64;
Query Match 4.48; Score 89; DB 1; Length 533;
Best Local Similarity 22.78; Pred. No. 4.5;
Matches 68; Conservative 42; Mismatches 77; Indels 112; Gaps 17;
OY 51 LMLTEERRNARELVSSEELGRMDALVVMGSDGLMEHYVGL-----MERPME 98
DB 28 LQVNRKQNLKRELLA-ELQDCEGLIVASATVTVADVINAARKLOVGRAGCVNDVLE 86
OY 99 TAIQPLCSLPAGSGNALAASLNHYAGYQVNTNEDLLTNCILLCRLLSPNNLLSLHTA 158
DB 87 AATRGVILVWNPNTNCSLSAA-----ELT-----CGMLMCIARQIDP----- 123
OY 159 SGLRLFSVLSLWGFIAVDVDESSEKRYRLGEMRTTLGTFLELAALRYRGR-LATLPVGR 217
DB 124 -----ATASMKDKWDR-----KKFMGTFL-----NGKTLGILGLGR 155
OY 218 VGSKTPA-----SPVV-----VQGGPYDAHLVPLEEPVSHMTVPVPPED 256
DB 156 IGRFAARMAQMGMTKTVGYDPLISPEVAASPCVQO-----LPLFE-----IWLPLCD 201
OY 257 FVLV-LALHSHLG---SEMFAPMGRCAAGVHLLFYVRAG-VSRAMLLRLFLAMEKGR 310
DB 202 FTVVHTPLPLSTGTGLNDSTFA---QCKKGVAVVNCARGIVDEGALLR---ALQSG 253
RESULT 15
SERA_HUMAN STANDARD: PRT: 533 AA.
ID SERA_HUMAN

AC 043175; Q9B001;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (3-PGDH).
GN PGDH OR PGDH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE-20179699; PubMed-10713460;
RA Cho H.M., Jun D.Y., Bae M.A., Ahn J.D., Kim Y.H.;
RT "Nucleotide sequence and differential expression of the human
RT 3-phosphoglycerate dehydrogenase gene.";
RL Gene 245:193-201(2000).
RN
RP SEQUENCE FROM N.A. AND VARIANTS PGDH DEFICIENCY MET-425 AND MET-490.
RX MEDLINE-20530221; PubMed-11053895;
RA Klomp L.W.J., de Koning T.J., Malingre H.E.M., van Beurden E.A.C.M.,
RA Brink M., Opdam F.L., Duran M., Jaeken J., Pineda M.,
RA van Malsdergen L., Poll-The B.T., van den Berg I.E.T., Berger R.;
RT "Molecular characterization of 3-phosphoglycerate dehydrogenase
RT deficiency - a neurometabolic disorder associated with reduced
RT L-serine biosynthesis.";
RL Am. J. Hum. Genet. 67:1389-1399(2000).
RN
RP
RP TISSUE-Brain, Lung, and Muscle;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DDBJ databases.
CC -1- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) -> 3-
CC 3-phosphohydroxypyruvate + NADH.
CC -1- PATHWAY: Serine biosynthesis; first step.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- DISEASE: Defects in PGDH are the cause of a deficiency
CC characterized by congenital microcephaly, psychomotor retardation,
CC and seizures.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF006043; AAB88664.1;
CC EMBL; AF171237; AAD51415.1;
CC EMBL; BC000303; AAH00303.1;
CC EMBL; BC001349; AAH01349.1;
CC EMBL; BC011262; AAH11262.1;
CC HSSP; P08328; 1PSD.
CC GeneW; HGNC:8923; PGDH.
DR
DR MIM: 606879;
DR
DR MIM: 601815;
DR
DR InterPro: IPR002162; D_2hyd-ac_dh.
DR Pfam: PF00389; 2-Hacid_dh; 1.
DR Pfam: PF02826; 2-Hacid_dh; 1.
DR PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2-HYDROXYACID_DH_3; 1.
DR Serine biosynthesis; Oxidoreductase; NAD; Disease mutation.
KM Serine biosynthesis; Oxidoreductase; NAD; Disease mutation.
FT ACT_SITE 226 236 SUBSTRATE-BINDING (BY SIMILARITY).
FT ACT_SITE 283 283 BY SIMILARITY.
FT ACT_SITE 425 425 V -> M (IN PGDH DEFICIENCY).
FT VARIANT 490 490 /FTID-VAR_013461.
FT VARIANT 490 490 V -> M (IN PGDH DEFICIENCY).
FT /FTID-VAR_013462.

FT CONFLICT 25 25 D -> E (IN REF. 1).
SQ SEQUENCE 533 AA; 56650 MW; C58EB72275C45B35 CRC64;
Query Match 4.3%; Score 87; DB 1; Length 533;
Best Local Similarity 23.1%; Pred. No. 6.7;
Matches 69; Conservative 42; Mismatches 76; Indels 112; Gaps 18;
QY 51 LMTERRRHARELRESEELGRMDALVWMSGDGLMEVYNGV-----MERPDWE 98
DQ 28 LQVEKQKLSKEELIA-ELQDCGLIVRSATVTVADVINAERKIQVGRAGTGVDVDE 86
QY 99 TAIQPLCSLPAGSGNALASLNHYAGYEQVTEDELLNCTLLCRLLSPMLLSLHFA 158
DQ 87 AATFRGLVMTTPGNSLSNA-----ELT-----CGMIMCLARQIFQ-----ATA 126
QY 159 SGLRLFSVLSLAWGFIADVDLESEKRYRGLGEMRTLCPLRLAALRTYRGR-LAYLPVGR 217
DQ 127 S-----MDGKMKRKF-----MGTEL-----NKTGLIGLGR 155
QY 218 VSKTTPA-----SPVY-----VOGPDVAHLVLEPVPVSHWTVPPDD 256
DQ 156 IGRVATRMOSFGMKRTIGYDPIISPEVSASFGVQD-----LPLEE-----IWPICD 201
QY 257 FVLV-LALLSHSLG---SEMFAPMGRCAGVHLEFVYRAG-VSRAMLRLFLAMEKGR 310
DQ 202 FITVHTPLPSTGTGLNDNTFA---QCKRGVRYVNCARGGIVDEGALLR---ALQSQ 253

Search completed: May 9, 2003, 16:55:19
Job time : 19 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:53:49 ; Search time 22 seconds
(without alignments)
1677.983 Million cell updates/sec

Title: US-09-937-060a-5

Perfect score: 2016

Sequence: 1 MDPAGGPGVLPFRCHVYL.....CVPPEPSMKPQMPPEEPL 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries
PIR_73:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	451	22.4	1240	2 T05162	hypothetical prote
2	374.5	18.6	473	2 T19707	hypothetical prote
3	356.5	17.7	687	2 S51398	hypothetical prote
4	355	17.6	458	2 T38776	hypothetical prote
5	345	17.1	624	2 S67059	hypothetical prote
6	250.5	12.4	549	2 T33517	hypothetical prote
7	155.5	7.7	306	2 AE1394	conserved hypotet
8	140.5	7.0	306	2 H69995	hypothetical prote
9	136.5	6.8	309	2 A83894	hypothetical prote
10	130	6.4	295	2 F69795	conserved hypotet
11	129.5	6.4	303	2 D86677	transcription regu
12	127	6.3	342	2 F86849	transcription regu
13	126	6.2	302	2 AG1665	hypothetical prote
14	124.5	6.2	310	2 F83871	multidrug resistan
15	122.5	6.1	311	2 F84898	hypothetical prote
16	122.5	6.1	364	2 A11293	hypothetical prote
17	119.5	5.9	309	2 AH1528	conserved hypotet
18	118	5.9	309	2 AB2166	hypothetical prote
19	117	5.8	315	2 AP1171	conserved hypotet
20	113	5.6	309	2 G95120	conserved hypotet
21	111	5.5	294	2 H70861	hypothetical prote
22	111	5.5	309	2 C97990	conserved hypotet
23	111	5.5	311	2 C70596	hypothetical prote
24	111	5.5	321	2 A89978	conserved hypotet
25	106	5.3	315	2 D83734	hypothetical prote
26	105.5	5.2	295	2 F72386	conserved hypotet
27	104	5.2	304	2 F69595	multidrug resistan
28	103	5.1	297	2 E72611	probable ATP-depen
29	103	5.1	1273	2	

30	100.5	5.0	433	2 S75948	hypothetical prote
31	99	4.9	449	2 C83634	hypothetical prote
32	97.5	4.8	305	2 G89844	hypothetical prote
33	97	4.8	311	2 D75405	conserved hypotet
34	94	4.7	899	2 T38153	gene retit protein
35	93.5	4.6	791	2 A46140	diacylglycerol kin
36	93.5	4.6	796	2 H46140	diacylglycerol kin
37	92.5	4.6	475	2 H84567	probable diacyglyl
38	92.5	4.6	732	2 T16422	hypothetical prote
39	92	4.6	782	2 T148746	semaphorin C - mou
40	92	4.6	1192	2 T17255	hypothetical prote
41	91.5	4.5	304	2 A86842	conserved hypotet
42	91.5	4.5	364	2 H87212	probable glucose e
43	91.5	4.5	1172	2 T36053	probable ABC-type
44	91	4.5	219	2 AE0724	probable hydrolyase
45	90.5	4.5	490	2 G85354	hypothetical prote

ALIGNMENTS

RESULT 1
T05162
hypothetical protein F18E5.160 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T05162
R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dijkse, W.; Stiekema, W.; Bancroft, I.;
submitted to the Protein Sequence Database, August 1998
A:Reference number: 215400
A:Accession: T05162
A:Molecule type: DNA
A:Residues: 1-1240 <BEV>
A:Cross-references: EMBL:AL022603
A:Experimental source: cultivar Columbia; BAC clone F18E5
C:Genetics:
A:Map position: 4
A:Introns: 44/3; 117/3; 178/3; 206/1; 237/2; 375/1; 416/1; 449/3; 474/1; 503/1; 529/1
A:Note: F18E5.160

Query Match 22.4%, Score 451, DB 2, Length 1240;

Best Local Similarity 31.9%, Pred. No. 1e-31;

Matches: 122; Conservative 63; Mismatches 127; Indels 70; Gaps 13;

QY	11	LPRPRVLYLNPGRGKKAQLFPSHQPLAEADISFTLMTERRNHRRLVSEELG	70
DB	374	LGRPRLLVFNVPFGSKSAREIFVKEVKPLFEDADVLEIQETVYQLHAKKFKSMDS	433
QY	71	RMDALVMSGDGLMHEVNVGLMERPDWETAIQKPLCSLPAGSGNALAASLNHYAGEQVT	130
DB	434	KYDGIKVSQDGLLEVNVNGLERADWRNALKPLIGMPAGNGMIMSLDVTGLRCA	493
QY	131	NEDLLTNCITLLICRRLSPMNLISHTASGURLFSVLSLWMTADVDLESKRYRLGEM	190
DB	494	N-----SATISITIRGHRKRSVDVATL-AOQNTKFFSVMLAMGLADIIDESEKPFMWSA	547
QY	191	FTTLGTFL-----RCLALRTYNGRLALYIPV-GRVSKTASPVVQV-	230
DB	548	RIDFVYVCLVDKFDNYCIAVVKLALQIRITICLRNRYGRILIFAPGFEQGGQASCLVOE	607
QY	231	-----OGP-----VDALHLPLEEPVSHMTVPDEDFVLVALLHS-HLGSE-MFA	274
DB	608	PHVSKKEYGOGPEKTFKFDLEKREMKGFPVTW-----LHNWMSGSENTLT	653
QY	275	APMGRCAGVMHLEFVRAGVSRAMLRLFLANEKGRHMEYECPLYLVY-----PVAFRL	329
DB	654	APAAKFSQGYLDLIVAK-NCPKLVLISLMRQSSGTHV--ESPPIYVYIKLVEKKAFAVL	710
QY	330	EP-----KDKGVFAVVDGELM	345
DB	711	EPGALVDEPKDEGIIDSGEVL	732

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: A83894
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: A83894
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <STO>
A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA05672.1; GSPDB:GNOC
C:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1953

Query Match 6.4%; Score 130; DB 2; Length 295;
Best Local Similarity 22.0%; Pred. No. 0.00076;
Matches 78; Conservative 50; Mismatches 138; Indels 88; Gaps 14;

QY 19 VLINPRGKRALQLFRRSHVOPPLAEAFISFTLMTERRNHARELVRSSEELGRMDALVVM 78
D 4 L I V K A S G N G K G R T W K K - V E Y E L Q I R N T P Y L V R F T S G S G A T T I V K L L E G V K T I I A V 62
QY 79 SGDLMEHVYVGLMERPDWETAIOKPLCSLPAGSGNALAASLNHYAGTEQYTNEDLTNC 138
D 63 GGDGTINEVANGLVNHH-----RVPGLIIPAGSGNDPACRLNTPMHYEKALHR----- 109
QY 139 TLLLCRRLSPMLSLHTASGLRFSVLSLAWGFADV--DLESEKRR-----LGE 190
D 110 ---LFEKKOKKVVLDLHGLQRCILVTGT-----GFDKIAKTIVTEAIKNNFNQGFEGCL 161
QY 191 RFTLGTFLRLALRTYRGRALVLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPVSHMT 250
D 162 SYVLS---MLEVLDKDYR-----PTNIQIT 182
QY 251 VVPDEDF---VLVIALHLS-HLGSEMFAPKRCAGVHLEFYRAGVSRMLRLFLIAM 306
D 183 VDGELEFSGVWLVAVANSPVYGGGIRICPEASVDDGLNICVHA-GMSKQMLRLFLPKA 241
QY 307 EKGRR--MEYECPLY-----VVPVVAFLRLEPKDKGFAVAGELVSSANVOGV 354
D 242 YKGRHVMEQHVTLTGKDYV-----OSDTFVVLVOSDGEPIESPRLOI 287

RESULT 11
F69795
conserved hypothetical protein yero - *Bacillus subtilis*
C:Species: *Bacillus subtilis*
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: F69795
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capriano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koehler, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, S.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadleir, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeder, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Serot
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, J.; Tosiato, V.; Uchiyama
T.; Winkler, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zunsstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:96044033; PMID:9384377
A:Accession: F69795
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-303 <KUN>
A:Cross-references: GB:299107; GB:AL009126; NID:g2632866; PIDN:CAB12492.1; PID:ell182652;
A:Experimental source: strain 168
C:Genetics:

A:Gene: yero

Query Match 6.4%; Score 129.5; DB 2; Length 303;
Best Local Similarity 20.6%; Pred. No. 0.00087;
Matches 68; Conservative 44; Mismatches 113; Indels 105; Gaps 13;

QY 16 RVVLINPRGKRALQLFRRSHVOPPL--AEAFISFTLMTERRNHARELVRSSEELGRW 72
D 3 RARILYNPTSGR---EIFKKHQAQVLOKFEQAGYETSHATTCAGDTHAAKEALREF 58
QY 73 DALVMSDDGLMEHVYVGLM---ERPDETAIOKPLCSLPAGSGNALAASLNHYAGTEQY 129
D 59 DLIIAAGDGTINEVANGLVNHH-----TLGVIPIVGTINDFARALG----- 103
QY 130 TNEDLTNCITLLLCRRLSPMLN-----LSLHTASGLRFSVLSLAWGFADVPLESEK 183
D 104 PREDILKADIVY-NGVARPIDIGVNGQYFNTIAGGRLTEL-----TYVPSKL 153
QY 184 YRLGEMFTLOTFLRLAALR-----TYRGRALVLPVGRVSKTPASPVVVOGQPVDAH 237
D 154 KTMGLQALAYLKGMELPSLRTEVEIEYDKLF-----QGEIMLF 194
QY 238 LVPLEEPVSHMTVVPDEDF-----VLVIALHLSHGSEMFAPM 277
D 195 LVTITNSVGEFEKLPADSSLDNMDMLIKKANLAFTIRATMALRGEHINDQ----- 248
QY 278 GRCAGVHMLFY-----VRAGVSRMLRL 302
D 249 -----HIITKANRYKANVSEKQMLNL 270

RESULT 12
D86677
transcription regulator [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)
C:Species: *Lactococcus lactis* subsp. *lactis*
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: D86677
R:Bojotin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Eh
genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis*
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: D86677
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-342 <STO>
A:Cross-references: GB:AE005176; PID:g12723295; PIDN:AAK04518.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yecE

Query Match 6.3%; Score 127; DB 2; Length 342;
Best Local Similarity 20.4%; Pred. No. 0.0017;
Matches 72; Conservative 55; Mismatches 122; Indels 104; Gaps 15;

QY 19 VLINPRGKRALQLFRRSHVOPPLAEE---EISFTLMTERRNHARELVRSSEELGRMD 73
D 6 L I V N P T S G Q ---E I K K Y T A D I L D K L E Q Y E A S A Y Q T T A Q D S A K K A A R A T E A G - F D 60
QY 74 ALVMSGDLMEHVYVGLM---ERPDETAIOKPLCSLPAGSGNALAASLNHYAGTEQY 130
D 61 LIIAAGDGTINEVANGVSPFEKRPD-----LAIYPTGTINDFARALKIPRG----- 107
QY 131 NEDDLTNCITLLLCRRLSPMLNLSLHTASGLRFSVLSLAWGFADVPLESEKRYRIGEM 190
D 108 -----KPLEAIEI-----IKKNOITLNDVGHAVIREFODEY----- 139
QY 191 RFTLGTFLRLAN-----LRTYRGRALVLPVGRVSKTPASPVVVOGQPVDA 236
D 140 -----FIIIAAGGGLTETLVSPSHLKTAPFLAYLAKG-----A 174
QY 237 HLVPLEEPVSHMTVVPDE-----DFVLVIALHLSHGSEMFAPKRCAGVHMLFYR 291
D 175 ELLPVRKAPVR--VVHGEVPEGDISMFFALITNSVGEFEKFIADAKLDLGLFTLILVK 232

OY 292 AGVSRAMLRLFLAMEGRHMEYECPLYVVPVAFRLPEPKDGKGFVA-VDGE 343
 Db 233 TD-NLELIALALAIYANGKHL--DDVNLLEYIKTSKIELEALOGKILINDGE 282

RESULT 13

F86849
 C:Accession: F86849
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 R:Biologist: A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp.
 A:Reference number: AB6625; MUID:21235186; PMID:11337471
 A:Accession: F86849
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-302 <STO>
 A:Cross-references: GB:AE005176; PID:g12724823; PIDN:AAK05896.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:

Query Match 6.2%; Score 126; DB 2; Length 302;
 Best Local Similarity 21.3%; Pred. No. 0.0018;
 Matches 75; Conservative 64; Mismatches 123; Indels 92; Gaps 17;
 OY 19 VLLNPRGGKALQFRSHVOPPLAEAFISFTMLTERNNHARELYRS--ELGRMDALV 76
 Db 5 LLANPNSGAGKARLETLLT-PYLEKNNEYRLFKTAAGECALRQILDKNPDHVL 63
 OY 77 VMSGDGLMEYVNGMLERPDWETAIOKPLCSLPAGSNMLASL-----NHVAGYEQ 128
 Db 64 ILGGDGLTSLVINELEP-----EENAFSTIPSGSNDGFARSILKIKLDPLESEPARRG 115
 OY 129 VTNEDLTNCTLLCRRLSPNNLSLHTASGLRSLVSLSLAFVADYDESEKYERLG 168
 Db 116 INHETFLMN-----YQSGLSGIVLNIGIGDALT-VKSNNECKLK 156
 OY 189 EM-RFLGTFELALARTYRGRALYLPVGVSGSKTPASVYVVOGPDVAILVPLEEVP 246
 Db 157 QVLMKLLGSE-----SYL-----TLALHVLITKRPFA-LIEVEN--- 191
 OY 247 SHMTVYDEDFVYLALLSHLSGSEMPFAWRCAAGVMLFVYRAGVSRAMLRLFL- 304
 Db 192 ---QEISLENAFLMTTKRPHFGCGVKISPEATMENADITL---VEYKHHLITRTSLI 244
 OY 305 -AMKGRHMEYECPLYVVPVAFRLPEPKDGKGFVAVDGLAVSEAVO--GGVH 355
 Db 245 PSYLMKHLNH-----PLPLAHVRSQ-----FSV--ELAESQVVDIGELH 283

RESULT 14

AC1665
 C:Accession: AC1665
 C:Species: Listeria innocua
 C:Date: 27-Nov-2003 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 R:Biologist: F.; Fringuel, L.; Buchrieser, C.; Amend, A.; Bagnaro, F.; Berche, P.; Bloeker
 K.; Glaser, F.; Hingault, G.; Duchaud, E.; Durand, L.; Dussutge, O.; Entian, K.D.; Fstli, H.
 J.; Dominguez Vernal, G.; Duchaud, E.; Durand, L.; Dussutge, O.; Entian, K.D.; Fstli, H.
 D.; Jones, L.; Kuhn, M.; Kunst, F.; Kurpkat, G.; Madueno, E.; Maitounnam, A.; Me
 A.; O.; Saluette, J.; Sines, N.; Tioroz, A.; Vazquez Boland, J.A.; Voss, H.; Weiland,
 A.; Filee, Compulsive of Listeria species
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AC1665
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-310 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC97095.1; PID:g16414366; GSPDB:GN00178

A:Experimental source: strain C1ip11262
 C:Genetics:
 A:Gene: 11n1865

Query Match 6.2%; Score 124.5; DB 2; Length 310;
 Best Local Similarity 21.3%; Pred. No. 0.0025;
 Matches 74; Conservative 55; Mismatches 118; Indels 101; Gaps 15;
 OY 19 VLLNPRGGKALQFRSHVOPPLAEAFISFTMLTERNNHARELYRS--ELGRMDALV 76
 Db 7 VLYNPTSIR-----EIKKNLADVLSTLEBQGVTSAMHATPEPDAAHAAEEAVRO---- 58
 OY 71 RMDALVMSGDLMEYVNGMLERPDWETAIOKPLCSLPAGSNMLASLNHVAGYEQT 130
 Db 59 RYDLYVAGCGGTINEVINGIADP-----YRPKGIIPGTITNDA----- 100
 OY 131 NEDLTNCTLLCRRLSPNNLSLHTASGLRSLVSLSLAFVADYDESEKYE----- 184
 Db 101 -----RALHVRDV-----IKATKIIAAGSVAMDIKANDOTYFINIGC 139
 OY 185 -RLGEMFTLCTFLRLAALRTYRGRALY-----LPVGRVSGSKTPASVYVVOGPDVA 236
 Db 140 GGRLEFLVYDP-----SRKTMIGOLATYTLKGIEMLP-----SLKATKVVVEYDQ- 186
 OY 237 -HLVPLEEFPVSHMTVYDEDFVYLALLSHLSGSEMPFAWRCAAGVMLFVYRAGVSR 296
 Db 187 -----VFEGVWFLLGLTNSIGGFKEI-APDAKLDDCKFSLIIVK-VNL 230
 OY 297 AMMLRLFLAMEGRHMEYECPLYVVPVAFRLPEPKDGKGFVAVDGL 344
 Db 231 AEFIRLVTLALRGDIKE--PNVIYVSEKSVSHSD-KMLINDGEL 275

RESULT 15

F83871
 C:Accession: F83871
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 R:Biologist: H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
 A:Reference number: AB3650; MUID:20512582; PMID:11058132
 A:Accession: F83871
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-311 <STO>
 A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BAR05493.1; GSPDB:G
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: bmrU

Query Match 6.1%; Score 122.5; DB 2; Length 311;
 Best Local Similarity 30.5%; Pred. No. 0.0038;
 Matches 32; Conservative 17; Mismatches 47; Indels 9; Gaps 3;
 OY 16 RVVLLNPRGGKALQFRSHVOPPLAEAFISFTMLTERNNHARELYSEELGRDAL 75
 Db 4 RAVLLVNCAGAGKGT-DKYLKDVPRALQSIGELTTPSKEGDIESICERREV---DLI 59
 OY 76 VVMSGDGLMEYVNGMLERPDWETAIOKPLCSLPAGSNMLASL 120
 Db 60 IVMGCGGTVECTINGLAVSPS-----PPLVVLPTGTCTNDPFRSL 99

Search completed: May 9, 2003, 16:57:23
 Job time: 26 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 03:43:39 ; Search time 4365 Seconds
(without alignments)
10487.678 Million cell updates/sec

Title: US-09-937-060A-19

Perfect score: 1573
Sequence: 1 gccccacagccgcgcctgcg.....gcctgaaaaaaaaaaaaa 1573

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sy:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_on:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_inv:*
33: em_htg_other:*
34: em_htg_mus:*
35: em_htg_pln:*
36: em_htg_rod:*
37: em_htg_mam:*
38: em_htg_vrt:*
39: em_sy:*
40: em_htgo_hum:*
41: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1550.8	98.6	1783	9 AF238083	AF238083 Homo sapi
2	1549.2	98.5	1824	9 BC008040	BC008040 Homo sapi
3	1547	98.3	1821	9 AK022402	AK022402 Homo sapi
4	1539.4	97.9	1772	9 AK023393	AK023393 Homo sapi
5	1539	97.8	1719	6 AX127641	AX127641 Sequence
6	1527	97.1	1693	9 AF26756	AF26756 Homo sapi
7	1511.6	96.1	1600	6 AX224379	AX224379 Sequence
8	1506.8	95.8	1869	9 BC009419	BC009419 Homo sapi
9	1416.4	90.0	2502	9 AK095578	AK095578 Homo sapi
10	1385.6	88.1	1428	9 HSA245504	AJ245504 Homo sapi
11	1306.6	83.1	2015	9 AB046025	AB046025 Macaca fa
12	1165	74.1	1173	9 AF200328	AF200328 Homo sapi
13	1155	73.4	1155	6 AX127642	AX127642 Sequence
14	1153.4	73.3	1192	9 BC004112	BC004112 Homo sapi
15	1052.8	66.9	183443	2 AC068145	AC068145 Homo sapi
16	1044	66.4	209861	2 AC021196	AC021196 Homo sapi
17	853.4	54.3	1815	6 AX287137	AX287137 Sequence
18	853.4	54.3	1815	10 AF068748	AF068748 Mus muscu
19	853.2	54.2	1759	6 AX224381	AX224381 Sequence
20	824.8	52.4	1842	10 AB049573	AB049573 Rattus no
21	824.8	52.4	1895	10 AB049572	AB049572 Rattus no
22	824.8	52.4	1943	10 AB049571	AB049571 Rattus no
23	824.8	52.4	2501	10 AB049574	AB049574 Rattus no
24	824.8	52.4	2648	10 AB049575	AB049575 Rattus no
25	806.2	51.3	1559	6 AX287138	AX287138 Sequence
26	806.2	51.3	1559	10 AF068749	AF068749 Mus muscu
27	763.2	48.5	1146	10 AF415213	AF415213 Mus muscu
28	541.6	34.4	110000	2 AL606505	AL606505 Mus muscu
29	541.6	34.4	112037	10 AL645851	AL645851 Mouse DNA
30	387.6	24.6	394	6 AX334274	AX334274 Sequence
31	262.8	16.7	2380	9 AF245447	AF245447 Homo sapi
32	262.8	16.7	2731	9 BC010671	BC010671 Homo sapi
33	262.8	16.7	2875	6 AX086301	AX086301 Sequence
34	262.8	16.7	2875	9 HSM801669	AL136701 Homo sapi
35	262.8	16.7	3012	9 BC006161	BC006161 Homo sapi
36	250	15.9	1854	10 AF415214	AF415214 Mus muscu
37	250	15.9	2416	10 BC006941	BC006941 Mus muscu
38	246.8	15.7	2698	10 AF245448	AF245448 Mus muscu
39	240	15.3	240	6 AX127662	AX127662 Sequence
40	193	12.3	296	6 AX198252	AX198252 Sequence
41	193	12.3	296	6 AX208818	AX208818 Sequence
42	178.4	11.3	44767	2 AC106591	AC106591 Rattus no
43	177.8	11.3	254729	2 AC122611	AC122611 Rattus no
44	159.2	10.1	1172	9 AK000599	AK000599 Homo sapi
45	138	8.8	173324	2 AC109120	AC109120 Rattus no

ALIGNMENTS

RESULT 1
AF238083
LOCUS AF238083 1783 bp mRNA linear PRI 01-JUN-2000
DEFINITION Homo sapiens sphingosine kinase-1 mRNA, complete cds.
ACCESSION AF238083
VERSION AF238083.1 GI:8132867
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1783)
Nava,V.E., Lacana,E., Poulton,S., Liu,H., Sugiyama,M., Kono,K.,
Milstien,S., Kohama,T. and Spiegel,S.
TITLE Functional characterization of human sphingosine kinase-1

JOURNAL FEBS Lett. 473 (1), 81-84 (2000)
MEDLINE 20263733
PUBMED 10802064
REFERENCE 2 (bases 1 to 1783)
AUTHORS Nava, V. E., Lacana, E., Poulton, S., Liu, H., Suglione, M., Kono, K.,
Milstien, S., Kohama, T. and Spiegel, S.
TITLE Direct Submission
JOURNAL Submitted (23 FEB-2000) Biochemistry, Georgetown University, 3900
Reservoir Rd NW BSB km 357, Washington, DC 20007, USA
FEATURES
Source Location/Qualifiers
1..1783
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q25.2"
337..1491
/codon_start=1
/product="sphingosine kinase-1"
/protein_id="AF73423.1"
/db_xref="GI:8132868"
/translation="MDPAGPRGVLPKPRVLLINPRGGKALQFHSVHOPILAE
AEISFTMLTERNHARELYSEELRWDALVWMSGDLNHEVYVNGLMERPMETALQ
KPLCSLPAGSNALASLNHYAGEOVNEDLLNCTLLCRILSLPMLLSLHPSG
LRLPSVLAMGFIADVDLESEKRYRLEMEFTLTGLRLALRLRYRGLAYLPGRV
GSKPASPVVQGGADVAHVPLEPVPVSHMTWPPDEPFLIILALSHSHEMERAP
MRCARAGVMLFYPVAGVSRAMLRLPFLAMGHEMPECPYLVYVPAVFLERKDC
KCFVAVDGEMLSEYAVQGVHNPYFMVSGCVPEPPSMKFOQMPPEEPL"
BASE COUNT 314 a 546 c 591 g 332 t
ORIGIN
Query Match 98.68; Score 1550.8; DB 9; Length 1783;
Best Local Similarity 99.58; Pred. No. 7.1e-219;
Matches 1566; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 1 GCCCCACAGCCGCCCTCCAGCCGCCCTGGGACACCCAGATGAAGAGTGAAGCAGC 60
DB 207 GCCCCACAGCCGCCCTCCAGCCGCCCTGGGACACCCAGATGAAGAGTGAAGCAGC 266
QY 61 AGCCGCCGCCAGGGGACGCG-CCCCACAGCCGCCAGGAGACCCCTGGCAGCGGAGCGCG 119
DB 267 AGCCGCCGCCAGGGGACGCGCCGCCACAGCGCCAGGAGACCCCTGGCAGCGGAGCGCG 326
QY 120 GGTGAGGTTATGATTCAGCCGCGGCGGCCCGGCGTGTCCCGCGGCCCTCCGCGCT 179
DB 327 GGTGAGGTTATGATTCAGCCGCGGCGGCCCGGCGTGTCCCGCGGCCCTCCGCGCT 386
QY 180 GGTGAGGTTATGATTCAGCCGCGGCGGCCCGGCGTGTCCCGCGGCCCTCCGCGCT 239
DB 387 GGTGAGGTTATGATTCAGCCGCGGCGGCCCGGCGTGTCCCGCGGCCCTCCGCGCT 446
QY 240 GGTGAGGTTATGATTCAGCCGCGGCGGCCCGGCGTGTCCCGCGGCCCTCCGCGCT 299
DB 447 GGTGAGGTTATGATTCAGCCGCGGCGGCCCGGCGTGTCCCGCGGCCCTCCGCGCT 506
QY 300 GAACACAGCGCGGAGCTGTGCGGTGGAGAGCTGGCGCGCTGGAGACCTGTGTGT 359
DB 507 GAACACAGCGCGGAGCTGTGCGGTGGAGAGCTGGCGCGCTGGAGACCTGTGTGT 566
QY 360 CATGCTGGAGAGGGGCTGATGACAGAGTGTGTGAACGGGCTCATGAGCGGCGCTGACT 419
DB 567 CATGCTGGAGAGGGGCTGATGACAGAGTGTGTGAACGGGCTCATGAGCGGCGCTGACT 626
QY 420 GGAGACCGCATCCAGAGCCCTGTGTAGCTCCAGCAGGCTGTGACAGCGGCTGGC 479
DB 627 GGAGACCGCATCCAGAGCCCTGTGTAGCTCCAGCAGGCTGTGACAGCGGCTGGC 686
QY 480 AGCTTCCTGAACCATTTATGCTGTGTATGAGCAGGTGACCAATGAAGACCTCCAGACCA 539
DB 687 AGCTTCCTGAACCATTTATGCTGTGTATGAGCAGGTGACCAATGAAGACCTCCAGACCA 746
QY 540 CTGGAGAGCTATTGCTGTGCGCGCGGCTGCTGTACCATGAACCTGCTGTCTGCACAC 599
DB 747 CTGGAGAGCTATTGCTGTGCGCGCGGCTGCTGTACCATGAACCTGCTGTCTGCACAC 806

QY 600 GCGTTCGCGGCTGGCGCCCTTCCTGTCGTCACGCCCTGGGCTTCATGCTGATGT 659
DB 807 GCGTTCGCGGCTGGCGCCCTTCCTGTCGTCACGCCCTGGGCTTCATGCTGATGT 866
QY 660 GGACCTAGAGAGTGAAGATATGCGGCTGGGGAGATGCGCTTCACCTGGGACCTT 719
DB 867 GGACCTAGAGAGTGAAGATATGCGGCTGGGGAGATGCGCTTCACCTGGGACCTT 926
QY 720 CCTGCGTGTGACGCGCTTCGACACCTACCGCGCGGAGCTGGCTACCTCCCTAGGAAG 779
DB 927 CCTGCGTGTGACGCGCTTCGACACCTACCGCGCGGAGCTGGCTACCTCCCTAGGAAG 986
QY 780 ACTGGGTTCCAAAGACACCTGCTCCCGGTTGTGTGTCCAGAGAGGCCGCTAGATGCACA 839
DB 987 ACTGGGTTCCAAAGACACCTGCTCCCGGTTGTGTGTCCAGAGAGGCCGCTAGATGCACA 1046
QY 840 CCTGTGCGACCTGGAGAGCGCAGTGGCCCTCTCAGCTGAGCAGTGGTGGCCGACGAGACTT 899
DB 1047 CCTGTGCGACCTGGAGAGCGCAGTGGCCCTCTCAGCTGAGCAGTGGTGGCCGACGAGACTT 1106
QY 900 TGTGCTATTCCTGGCAGCTGCTGACATGCGACCTGGGAGTGAATGTTCTGACCCAT 959
DB 1107 TGTGCTATTCCTGGCAGCTGCTGACATGCGACCTGGGAGTGAATGTTCTGACCCAT 1166
QY 960 GGGCGGCTGTGCACTGGCTGATGATGCTGTTCTAGCTGTGGCGGAGACTGTCTGTC 1019
DB 1167 GGGCGGCTGTGCACTGGCTGATGATGCTGTTCTAGCTGTGGCGGAGACTGTCTGTC 1226
QY 1020 CATGCTGTCGCGCTCTTCCTGTCGTCATGAGAGAGGCGCATATGAGATATGATGCC 1079
DB 1227 CATGCTGTCGCGCTCTTCCTGTCGTCATGAGAGAGGCGCATATGAGATATGATGCC 1286
QY 1080 CTACTTGTATATGTGCGCGCTGTCGCTTCGCTTGGAGGCCCAAGGATGGAAAGTGT 1139
DB 1287 CTACTTGTATATGTGCGCGCTGTCGCTTCGCTTGGAGGCCCAAGGATGGAAAGTGT 1346
QY 1140 GTTTCAGATGATGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
DB 1347 GTTTCAGATGATGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1406
QY 1200 CTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
DB 1407 CTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1466
QY 1260 GGCAGCGCAGAAAGAGCCCTTATGAGCCCTGGGCGCGCTGTGCTTATGATGATGATGAT 1319
DB 1467 GGCAGCGCAGAAAGAGCCCTTATGAGCCCTGGGCGCGCTGTGCTTATGATGATGATGAT 1526
QY 1320 AGGACCTTCCT 1379
DB 1527 AGGACCTTCCT 1586
QY 1380 GAGACTCTCTGAGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 1439
DB 1587 GAGACTCTCTGAGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 1646
QY 1440 TGAAGTCTTGGGTGAGAGCCAGCTGAGTGGGCCAGCTGCTATGATTAAGGCTTCTAG 1499
DB 1647 TGAAGTCTTGGGTGAGAGCCAGCTGAGTGGGCCAGCTGCTATGATTAAGGCTTCTAG 1706
QY 1500 TTTGTTCTGAGACCCCGACCCCGACGAAACCAATCAATAAAGTGAATTCAGAGCTGA 1559
DB 1707 TTTGTTCTGAGACCCCGACCCCGACGAAACCAATCAATAAAGTGAATTCAGAGCTGA 1766
QY 1560 AAAAAAAAAAAAAA 1573
DB 1767 AAAAAAAAAAAAAA 1780
RESULT 2
BC008040
LOCUS BC008040 1824 bp mRNA linear PRI 12-JUL-2001

	DEFINITION	Homo sapiens, clone MGC:15041 IMAGE:3831657, mRNA, complete cds.
	ACCESSION	BC008040
	VERSION	BC008040.1 GI:14165485
	KEYWORDS	MGC.
	SOURCE	Homo sapiens.
	ORGANISM	Homo sapiens.
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
	AUTHORS	1 (bases 1 to 1824)
	TITLE	Strausberg, R.
	JOURNAL	Direct Submission Submitted (21-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2530, USA
	REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
	COMMENT	Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: ATCC/DCMP/DRP cDNA Library Preparation: Rubin Laboratory DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbioology.org contact: amadan@systemsbioology.org Anup Madan, Rachel Dickhoff, Jessica Fahay, Stephanie Ford, Julia Greene, Mark Keltman and Anuradha Madan
	FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 24 Row: n Column: 4 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10433790. Location/Qualifiers 1..1824 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="MGC:15041 IMAGE:3831657" /tissue_type="Skin, melanotic melanoma." /clone_id="NH_MGC_20" /lab_host="DH10B-R" /note="Vector: pOTB7" 339..1493 /codon_start=1 /product="Unknown (protein for MGC:15041)" /protein_id="AAH08040.1" /db_xref="gi:14165486" /translation="MDPAGGPGVLTAPRCRVILNLRGGKGAOLFRSHVOPLAE AELFTMLTERHNARELYRSEBELGMKDVLVMSGGLMHVEYVNGIMRPDWTATQ KPLDSLPAGSGNALAASLNHYAGVEQVTNDELNTCLDLICRIILSPMNLSLTPASG LRTFESVSLAMGEITADVDLESEKRYRIGENKFILGTFLLAALFTYGRSLAYLPVGHV GSKTFPSLVVVOOGPVDAHEVPLEEPYSMTVYPDDFYLALFTYLSHGISEMFADP MGRCAAVMHLFEYRAGVSRAMLRLRLAMEKGHHMYECOPYLVVVVAVRLEPKGS KGIVAVDGELMVSAVNOGVNHPRTFMWVSGCVERPPSMKFOQNPPEEPL"
	CDS	/codon_start=1 /product="Unknown (protein for MGC:15041)" /protein_id="AAH08040.1" /db_xref="gi:14165486" /translation="MDPAGGPGVLTAPRCRVILNLRGGKGAOLFRSHVOPLAE AELFTMLTERHNARELYRSEBELGMKDVLVMSGGLMHVEYVNGIMRPDWTATQ KPLDSLPAGSGNALAASLNHYAGVEQVTNDELNTCLDLICRIILSPMNLSLTPASG LRTFESVSLAMGEITADVDLESEKRYRIGENKFILGTFLLAALFTYGRSLAYLPVGHV GSKTFPSLVVVOOGPVDAHEVPLEEPYSMTVYPDDFYLALFTYLSHGISEMFADP MGRCAAVMHLFEYRAGVSRAMLRLRLAMEKGHHMYECOPYLVVVVAVRLEPKGS KGIVAVDGELMVSAVNOGVNHPRTFMWVSGCVERPPSMKFOQNPPEEPL"
	BASE COUNT	354 a 351 c 592 g 327 t
	ORIGIN	
	Query Match	98.5% Score 1549.2 DB 9 Length 1824;
	Best Local Similarity	99.4% Pred No. 1.4e-278;
	Matches 1565:	Conservative 0 Mismatches 8 Indels 1 Gaps 1;
0Y	1	GCCCCACAGCGGCCCTTGCGACGCGCCCTGCGGCGACGCGATTAAGAGCTGAAGCGCAGG 60
Db	209	GCCCCACAGCGGCCCTTGCGACGCGCCCTGCGGCGACGCGATTAAGAGCTGAAGCGCAGG 268
0Y	61	AGCGCGCGCCACAGGCGCAGCG-CCCCACAGCGCGCACAGGCCCTTGCGACAGCGGAGCGCG 119
Db	269	AGCGCGCGCCACAGGCGCAGCGCGCCCGCCACAGCGCGCACAGGCCCTTGCGACAGCGGAGCGCG 328
0Y	120	GGTGAGGTTTATGATTCACAGGCGGCGCGCGCGCGCGCGTGTCTCCCGCGCGCGT 179
Db	329	GGTGAGGTTTATGATTCACAGGCGGCGCGCGCGCGCGCGCGTGTCTCCCGCGCGT 388
0Y	180	GCTGCTGCTGCTGAACCCGCGCGCGCGGACAAGGCGCATTCGAGCTCTTCGCGAGTCA 239

[illegible]

QY 900 TGTGCTAGTCCCTGCGACCTGCTGCACTGCGACCTGGGCACTGAGATGTTGCTGCACTGCACT 959
 DB 1163 TGTGCTAGTCCCTGCGACCTGCTGCACTGCGACCTGGGCACTGAGATGTTGCTGCACTGCACT 1222
 QY 960 GGGCGGCTGCGACCTGCGACCTGCTGCACTGCGACCTGGGCACTGAGATGTTGCTGCACTGCACT 1019
 DB 1223 GGGCGGCTGCGACCTGCGACCTGCTGCACTGCGACCTGGGCACTGAGATGTTGCTGCACTGCACT 1282
 QY 1020 CATGCTGCTGCGACCTGCTGCGACCTGCGACCTGGGCACTGAGATGTTGCTGCACTGCACT 1079
 DB 1283 CATGCTGCTGCGACCTGCTGCGACCTGCGACCTGGGCACTGAGATGTTGCTGCACTGCACT 1342
 QY 1080 CTACTGCTGATATGTCGCGACCTGCTGCGACCTGGGCACTGAGATGTTGCTGCACTGCACT 1139
 DB 1343 CTACTGCTGATATGTCGCGACCTGCTGCGACCTGGGCACTGAGATGTTGCTGCACTGCACT 1402
 QY 1140 GTTTCGATGATGTCGCGACCTGCTGCGACCTGGGCACTGAGATGTTGCTGCACTGCACT 1199
 DB 1403 GTTTCGATGATGTCGCGACCTGCTGCGACCTGGGCACTGAGATGTTGCTGCACTGCACT 1462
 QY 1200 CTACTGCTGATATGTCGCGACCTGCTGCGACCTGGGCACTGAGATGTTGCTGCACTGCACT 1259
 DB 1463 CTACTGCTGATATGTCGCGACCTGCTGCGACCTGGGCACTGAGATGTTGCTGCACTGCACT 1522
 QY 1260 GCGACCGCGACGAG 1319
 DB 1523 GCGACCGCGACGAG 1582
 QY 1320 AGGACCT 1379
 DB 1583 AGGACCT 1642
 QY 1380 GAGACT 1439
 DB 1643 GAGACT 1702
 QY 1440 TGAAGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1499
 DB 1703 TGAAGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1762
 QY 1500 TTTGTTCTGAGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCG 1558
 DB 1763 TTTGTTCTGAGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCCGACCCG 1821
 RESULT 4
 AK023393
 LOCUS Homo sapiens cDNA FLJ13331 fls, clone OVARC1001809, moderately similar to Mus musculus sphingosine kinase (SPHK1a) mRNA. AK023393
 ACCESSION AK023393.1 GI:10435311
 VERSION 1
 KEYWORDS oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens ovary, tumor tissue cDNA to mRNA, clone_11b.OVARC1 clone:OVARC1001809.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wagaetsuma,M., Hosokita,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Matanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuno,Y., Ninomiya,K. and Iwayanagi,T.
 TITLE NED0 human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1772)
 AUTHORS Isogai,T. and Otsuki,T.
 TITLE Direct Submission
 JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

COMMENT (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NED0 human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology: cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
 FEATURES
 source.
 1..1772
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="OVARC1001809"
 /tissue_type="ovary, tumor tissue"
 /clone_11b="OVARC1"
 /note="cloning vector: pME18FL3"
 350..1750
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB1458.1"
 /translation="MDPAGPRGVLPKRCVLYLNPQGGKKAQLQFRSHVPLAE
 AEISFTLTERNNHARELVSEELRMALVMSGDGLMHEVNGLMRPMPEALQ
 KPLCSLPAGSGNALAASLNHYAGYEQYMEDLLTNCITLLCRLLSPMLLSLHTASG
 LRLFSVLSLAWGFIADVDESEKRYLGRMFTLFTLALRLYRGLAYLPVGRV
 GSKTPASPVVGGPDAILVPLEPVPVHMVDPEDVFLVIALVLSHLSGDMFAAP
 MGRCAAGVHMLFYVRAGVRAALLRLFLAMEGRHREYCPILVYVPAFLERKDS
 KGVFAVDGLMSEAVQGVHPNTEWMSGVCEVPEPSMKPQMPPEEPL"
 BASE COUNT. 285 a 566 c 585 g 336 t
 ORIGIN
 Query Match 97.9%; Score 1539.4; DB 9; Length 1772;
 Best Local Similarity 99.9%; Pred. No. 9.4e-277;
 Matches 1551; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 GCCCCACAGCCGCGCTCTGCGACGCGCGCTGGGCGACGCGATAGAGCTGAAGCAGG 60
 DB 220 GCCCCACAGCGGCGCTCTGCGACGCGCGCTGGGCGACGCGATAGAGCTGAAGCAGG 279
 QY 61 AGCCGCGCGCAGGCGAGCGG -CCCCACAGCGCGCAGGAGCCCTCTGGCAGCGGAGCGCG 119
 DB 280 AGCCGCGCGCAGGCGAGCGGCGCGCCACAGCGCGCAGGAGCCCTCTGGCAGCGGAGCGCG 339
 QY 120 GGTGAGGTTATGATTCACGCGGCGCGCGCGCGCGCGCTCCCGGCGCTGCGCGCGT 179
 DB 340 GGTGAGGTTATGATTCACGCGGCGCGCGCGCGCGCGCGCTCCCGGCGCTGCGCGCGT 339
 QY 180 GCTGTCCTGCTGACCCGCGCGCGCGCGCGCAAGGAGCCTTCAGCTCTCCGAGTCA 239
 DB 400 GCTGTCCTGCTGACCCGCGCGCGCGCGCGCAAGGAGCCTTCAGCTCTCCGAGTCA 459
 QY 240 CGTGACAGCCCTTTTGGCTGAGCTGAATCTCTTACGCTGATGCTCACTAGCGGCG 299
 DB 460 CGTGACAGCCCTTTTGGCTGAGCTGAATCTCTTACGCTGATGCTCACTAGCGGCG 519
 QY 300 GAACCAAGCGCGGAGCTGTGCGTCGCGAGAGCTGGCGCTGGGAGCCTCTGGGCGT 359
 DB 520 GAACCAAGCGCGGAGCTGTGCGTCGCGAGAGCTGGCGCTGGGAGCCTCTGGGCGT 579
 QY 360 CATGCTGAGACGCGGCTGATGACAGAGTGTGAACGCGGCTCATGAGCGGCGTCACTG 419
 DB 580 CATGCTGAGACGCGGCTGATGACAGAGTGTGAACGCGGCTCATGAGCGGCGTCACTG 639
 QY 420 GAGAGCGCGCATGACAGAGCCCTGTGTACCTTCCAGAGAGCTCTGGCAACGCGTGGC 479
 DB 640 GAGAGCGCGCATGACAGAGCCCTGTGTACCTTCCAGAGAGCTCTGGCAACGCGTGGC 699
 QY 480 AGCTTCCTTAACCATTTATGCTGATGAGCGAGTGCACCAATGAAGACCTCTGACCAA 539
 DB 700 AGCTTCCTTAACCATTTATGCTGATGAGCGAGTGCACCAATGAAGACCTCTGACCAA 759
 QY 540 CTGACGCTATGCTGCGCGCGGCTGCTGTCACCATGACCATGCTGCTGTCACAC 599

Db 760 CTGCACGCTATTCCTGCTGCGCGGCGCTGCTGCACCCATGACCTCTGCTCTGCAACAC 819
OY 600 GGGCTTGG 659
Db 820 GGGCTTGG 879
OY 660 GGACCTAG 719
Db 880 GGACCTAG 939
OY 720 CCTGCGTCTGGCAGCCCTGCGACCTACCGCGCGCGAGCTGAGCTCTGCTGAGAG 779
Db 940 CCTGCGTCTGGCAGCCCTGCGACCTACCGCGCGCGAGCTGAGCTCTGCTGAGAG 999
OY 780 AGTGGGTTCCAGACACACCTGCTCCCGGTTGTGTCTCAGACAGGGCCCGGTAGATGACA 839
Db 1000 AGTGGGTTCCAGACACACCTGCTCCCGGTTGTGTCTCAGACAGGGCCCGGTAGATGACA 1059
OY 840 CCTGTGCGACCTGAG 899
Db 1060 CCTGTGCGACCTGAG 1119
OY 900 TGTGCTAGTCTGGCAGCTGCTGACCTGCGACCTGGGAGAGATGTTTGTGCTGACCCAT 959
Db 1120 TGTGCTAGTCTGGCAGCTGCTGACCTGCGACCTGGGAGAGATGTTTGTGCTGACCCAT 1179
OY 960 GGGCGCGTGGCAGCTGGGGGATGATCATCTGTCTAGCTGGGGGGAGTGTCTGCTGCG 1019
Db 1180 GGGCGCGTGGCAGCTGGGGGATGATCATCTGTCTAGCTGGGGGGAGTGTCTGCTGCG 1239
OY 1020 CATGCTGCTGCGCTCTTCTTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1079
Db 1240 CATGCTGCTGCGCTCTTCTTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1299
OY 1080 CTACTTGGTATATGTGCGCGGTGCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1139
Db 1300 CTACTTGGTATATGTGCGCGGTGCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1359
OY 1140 GTTTCAGAGTGGAG 1199
Db 1360 GTTTCAGAGTGGAG 1419
OY 1200 CTACTTGGTATATGTGAGCGGTTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1259
Db 1420 CTACTTGGTATATGTGAGCGGTTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1479
OY 1260 GCGACCGCCGAG 1319
Db 1480 GCGACCGCCGAG 1539
OY 1320 AGGACCTTCT 1379
Db 1540 AGGACCTTCT 1599
OY 1380 GAGACCT 1439
Db 1600 GAGACCT 1659
OY 1440 TGAAGTCT 1499
Db 1660 TGAAGTCT 1719
OY 1500 TTTGTCTGAG 1552
Db 1720 TTTGTCTGAG 1772

RESULT 5
AX127641 1719 bp DNA Linear PAT 15-MAY-2001
LOCUS AX127641
DEFINITION Sequence 1 from Patent W00131029.
ACCESSION AX127641
VERSION AX127641.1 GI:14134307

KEYWORDS
SOURCE human
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1719)
AUTHORS Allen, J., Gosink, M., Melendez, A.J. and Takacs, L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A1 03-May-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
source
1. 1719
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 305 a 529 c 556 g 329 t
ORIGIN

Query Match 97.8%; Score 1539; DB 6; Length 1719;
Best Local Similarity 99.0%; Pred. No. 1.1e-276;
Matches 1548; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 11 CGGCGCTGCGAGCGCCGCTGGGAG 70
Db 151 CGGCGCTGCGAGCGCCGCTGGGAG 210
OY 71 ACGGCGAGCGCCCGACAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 130
Db 211 ACGGCGAGCGCCCGACAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 270
OY 131 TGAATCAG 190
Db 271 TGAATCAG 330
OY 191 TGAATCAG 250
Db 331 TGAATCAG 390
OY 251 TTTTGGCTGAGGCTGAATCTCTTCTGAGCTGATGCTGATGAGAGAGAGAGAGAGAGAGAG 310
Db 391 TTTTGGCTGAGGCTGAATCTCTTCTGAGCTGATGCTGATGAGAGAGAGAGAGAGAGAGAG 450
OY 311 GGGAGCTGTGCGGTCGAG 370
Db 451 GGGAGCTGTGCGGTCGAG 510
OY 371 ACGGCGTGAATGACAG 430
Db 511 ACGGCGTGAATGACAG 570
OY 431 TCCAG 490
Db 571 TCCAG 630
OY 491 ACCATTATGCTGCTATGAT 550
Db 631 ACCATTATGCTGCTATGAT 690
OY 551 TCGTGTGCGCGCGGCTGCTGACCATGAACTGCTGTCTGTGACAGAGAGAGAGAGAGAGAG 610
Db 691 TCGTGTGCGCGCGGCTGCTGACCATGAACTGCTGTCTGTGACAGAGAGAGAGAGAGAGAG 750
OY 611 TCGGCTCTTCTCTGTGCTGACGCTGAGCTGGGGCTGCAATGCTGATGTGAGAGAGAGAG 670
Db 751 TCGGCTCTTCTCTGTGCTGACGCTGAGCTGGGGCTGCAATGCTGATGTGAGAGAGAGAGAG 810
OY 671 GTGAGAGATATGAG 730
Db 811 GTGAGAGATATGAG 870
OY 731 CAGCGCTGCGACCTACCGCGCGAG 790
Db 871 CAGCGCTGCGACCTACCGCGCGAG 930

QY 791 AGACACCTGCTCCCGCCGCTGTGGTCCAGCAGGCGCCGCTAGATGACACCTTTGGCCAC 850
| | | | |
Db 931 AGACACCTGCTCCCGCCGCTGTGGTCCAGCAGGCGCCGCTAGATGACACCTTTGGCCAC 990
| | | | |
QY 851 TGGAGAGCAGTGGCCCT 910
| | | | |
Db 991 TGGAGAGCAGTGGCCCT 1050
| | | | |
QY 911 TGGCAGCTGCTGACCTGCGACCTGGGAGTGAAGATGTTTGTCTCACCACCTATGGGCGCCGCTGTG 970
| | | | |
Db 1051 TGGCAGCTGCTGACCTGCGACCTGGGAGTGAAGATGTTTGTCTCACCACCTATGGGCGCCGCTGTG 1110
| | | | |
QY 971 CAGCTGGCGTATGATGATCTGTTCTACGTGCGGGCGGAGATGTTCTCTGTCATGCTGTCG 1030
| | | | |
Db 1111 CAGCTGGCGTATGATGATCTGTTCTACGTGCGGGCGGAGATGTTCTCTGTCATGCTGTCG 1170
| | | | |
QY 1031 GCGCTTCTGCGCATGGAAGAGGAGGAGATATGAGATATGATGATGATGATGATGATGATGATGAT 1090
| | | | |
Db 1171 GCGCTTCTGCGCATGGAAGAGGAGGAGATATGAGATATGATGATGATGATGATGATGATGATGAT 1230
| | | | |
QY 1091 ATGTGCGCGTGGTCCGCTTCCGCTTGGAGCCGAGGATGGAAGGATGTTGTTGTCAGTGG 1150
| | | | |
Db 1231 ATGTGCGCGTGGTCCGCTTCCGCTTGGAGCCGAGGATGGAAGGATGTTGTTGTCAGTGG 1290
| | | | |
QY 1151 ATGGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1210
| | | | |
Db 1291 ATGGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1350
| | | | |
QY 1211 TGGTAGGCGGTGGTGGAGAGCCCGCCGAGTGGAGAGCCCGAGAGATGATGATGATGATGATGATGAT 1270
| | | | |
Db 1351 TGGTAGGCGGTGGTGGAGAGCCCGCCGAGTGGAGAGCCCGAGAGATGATGATGATGATGATGATGAT 1410
| | | | |
QY 1271 AAGAGCCCTTATGACCCCTGGGCGCGGCTGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 1330
| | | | |
Db 1411 AAGAGCCCTTATGACCCCTGGGCGCGGCTGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 1470
| | | | |
QY 1331 TCCTTCCCTAGGGGTGACAGGGCGCTGTCACAGCTCTGTTGGGGGTGGAGAGACTCTCT 1390
| | | | |
Db 1471 TCCTTCCCTAGGGGTGACAGGGCGCTGTCACAGCTCTGTTGGGGGTGGAGAGACTCTCT 1530
| | | | |
QY 1391 GGAAGAGGATGAGAGGATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1450
| | | | |
Db 1531 GGAAGAGGATGAGAGGATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1590
| | | | |
QY 1451 GTGAGAG 1510
| | | | |
Db 1591 GTGAGAG 1650
| | | | |
QY 1511 ACCCCACCCCGACGAAATCAAAATGAATGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1570
| | | | |
Db 1651 ACCCCACCCCGACGAAATCAAAATGAATGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1710
| | | | |
QY 1571 AAA 1573
| | | | |
Db 1711 AAA 1713

RESULT 6
AF266756 1693 bp mRNA linear PRI 01-JUN-2000
DEFINITION Homo sapiens sphingosine kinase (SPHK1) mRNA, complete cds.
ACCESSION AF266756
VERSION AF266756.1 GI:8133099
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Melendez, A.J., Carlos-Dias, E., Gosink, M., Allen, J.M. and Takacs, L.
TITLE (bases 1 to 1693)
Human Sphingosine Kinase, Molecular Cloning, Functional
Characterization and Tissue Distribution
JOURNAL Gene (2000) In press

REFERENCE 2 (bases 1 to 1693)
AUTHORS Melendez, A.J., Carlos-Dias, E., Gosink, M., Allen, J.M. and Takacs, L.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-2000) Department of Molecular and Cellular
Biology, Jouveinal Parke-Davis, 11-13 rue de la Loge, Fresnes
94265, France
FEATURES
source
1..1693
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q25.2"
1..1693
/gene="SPHK1"
270..1424
/gene="SPHK1"
/codon_start=1
/product="sphingosine kinase"
/protein_id="AAF73470.1"
/db_xref="GI:8133100"
/translation="MDPAGGPRGVLPRRCRVLTLLNPRGKALOLFPRSHVPLAE
AISPFTLMTERRNHRRELVSSEELRMDALVYNSDGLMEHVNGLMERPMTALIG
KPLCSLPAAGSGNALAASLNHYACTFOVYNEDELITNCTLLCRLLSPNNLLSLHTASG
LRLESVLSLWGFADVDLESEKTRLEEMRFTLTFLAALRTYRGLAYLPVGRV
GSKTPASPVVVOQPDALVPLEDPVSHWVVDDEDEFTLVALLHSLSEMPAP
MGRCAGVNHLEFYRAGVSRAMLRLFLAMEGRIMECEPVLVYVPAVFAFLPKDG
KGFAVDELGLVSEAVSQVHPNRYFMVMSGCEVPPSPMKPQMPPEPEPL"
1676..1681
polyA_signal
/gene="SPHK1"
BASE COUNT 279 a 529 c 556 g 329 t
ORIGIN
Best Match 97.1% Score 1527; DB 9; Length 1693;
Query Local Similarity 99.4%; Pred No. 1.9e-274;
Matches 1333; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 11 CGGCCCTGCGAGCCCGCCTGGGCAACAGATAGAGCTGAAGGACAGAGCGCGCC 70
| | | | |
Db 151 CGGCCCTGCGAGCCCGCCTGGGCAACAGATAGAGCTGAAGGACAGAGCGCGCC 210
| | | | |
QY 71 ACGGCGAGCGCCCGACAGCGCCAGGAGCCCGCTGGCAGCGGAGCGCGGCTGACGTTA 130
| | | | |
Db 211 ACGGCGAGCGCCCGACAGCGCCAGGAGCCCGCTGGCAGCGGAGCGCGGCTGACGTTA 270
| | | | |
QY 131 TGGATCCAGCGGGGCGCCCGCGGGGCGTCCCGCGGCGCTGCGCGGCTGCTGTCG 190
| | | | |
Db 271 TGGATCCAGCGGGGCGCCCGCGGGGCGTCCCGCGGCGCTGCGCGGCTGCTGTCG 330
| | | | |
QY 191 TGAACCGCGCGCGGCGAGAGGCGCAAGGCGCTTGCAGCTCTCCGAGTCACTGACAGCCCG 250
| | | | |
Db 331 TGAACCGCGCGCGGCGAGAGGCGCAAGGCGCTTGCAGCTCTCCGAGTCACTGACAGCCCG 390
| | | | |
QY 251 TTTTGGCTGAGGCTGAATCTCTTACGCTGATGCTCACTGAGCGGCGAACCACGCGC 310
| | | | |
Db 391 TTTTGGCTGAGGCTGAATCTCTTACGCTGATGCTCACTGAGCGGCGAACCACGCGC 450
| | | | |
QY 311 GGGAGCTGTGGCGGTGGAGAGGAGCTGGGCGGCTGGAGAGCTGCTGGTGCATGCTGGAG 370
| | | | |
Db 451 GGGAGCTGTGGCGGTGGAGAGGAGCTGGGCGGCTGGAGAGCTGCTGGTGCATGCTGGAG 510
| | | | |
QY 371 ACGGCGTGAATGACGAGTGTGTGAACGGGCTCATGAGAGCGGCTGACTGAGAGACGCCCA 430
| | | | |
Db 511 ACGGCGTGAATGACGAGTGTGTGAACGGGCTCATGAGAGCGGCTGACTGAGAGACGCCCA 570
| | | | |
QY 431 TCCAGAGGCCCTGTGTACCTCCAGCAGGCTGTGGCAACGGCGTGGCAAGCTTCTTGA 490
| | | | |
Db 571 TCCAGAGGCCCTGTGTACCTCCAGCAGGCTGTGGCAACGGCGTGGCAAGCTTCTTGA 630
| | | | |
QY 491 ACCATTATGCTGGCTATGAGCAGGATGACCAATGAACATCCGAGCAACTGACGCTAT 550
| | | | |
Db 631 ACCATTATGCTGGCTATGAGCAGGATGACCAATGAACATCCGAGCAACTGACGCTAT 690
| | | | |
QY 551 TCGTGTGCCCGCGCGTGTGTACCCATGAACCTGCTGTCTGTGACAGCGGCTTGGGGC 610
| | | | |

|||||
Db 691 TCGCTGCCCGCGGCTGCTGCACCATGAACCTGCTGCTCTGCAACAGGCTTGCGGGC 750
Oy 611 TCGGCTCTTCTCTGCTGCTCAAGCTGGGCTTTCATTTGCTGATGTGACCTAGAGA 670
Db 751 TCGGCTCTTCTCTGCTGCTCAAGCTGGGCTTTCATTTGCTGATGTGACCTAGAGA 810
Oy 671 GTGAGAGATGCGGCTGCGGGAGATGCGCTTCACTCTGGGACCTTCTGCGCTGG 730
Db 811 GTGAGAGATGCGGCTGCGGGAGATGCGCTTCACTCTGGGACCTTCTGCGCTGG 870
Oy 731 CAGCCCTGGCAGCTACCGCGCGGCTGCTACCTCCCTGTAGAGAGATGGTTCCA 790
Db 871 CAGCCCTGGCAGCTACCGCGCGGCTGCTACCTCCCTGTAGAGAGATGGTTCCA 930
Oy 791 AGACACCTGCTCCCGCTTGTGTCTCAGCAGAGGCGGCTAGTATGACACCTTGTGCAC 850
Db 931 AGACACCTGCTCCCGCTTGTGTCTCAGCAGAGGCGGCTAGTATGACACCTTGTGCAC 990
Oy 851 TGGAGAGCAGTGGCTCTCTCACTGAGAGTGGCCGACGAGACTTGTGTCTGAGTCC 910
Db 991 TGGAGAGCAGTGGCTCTCTCACTGAGAGTGGCCGACGAGACTTGTGTCTGAGTCC 1050
Oy 911 TGGAGAGCAGTGGCTCTCTCACTGAGAGTGGCCGACGAGACTTGTGTCTGAGTCC 970
Db 1051 TGGAGAGCAGTGGCTCTCTCACTGAGAGTGGCCGACGAGACTTGTGTCTGAGTCC 1110
Oy 971 CAGCTGCGCTGATGATCTCTTCTACGTGGGCGGAGTGTCTGCTGCTGCTGCTGC 1030
Db 1111 CAGCTGCGCTGATGATCTCTTCTACGTGGGCGGAGTGTCTGCTGCTGCTGCTGC 1170
Oy 1031 GCTCTCTCTGCGCAGTGAAGAGGCGAGCATATGAGATGATGATGATGATGATGAT 1090
Db 1171 GCTCTCTCTGCGCAGTGAAGAGGCGAGCATATGAGATGATGATGATGATGATGAT 1230
Oy 1091 ATGTGCGCGTGTGCTGCTCTCTGCTGAGCGCCAAAGATGAGGAAAGTGTGTTGCACTG 1150
Db 1231 ATGTGCGCGTGTGCTGCTCTCTGCTGAGCGCCAAAGATGAGGAAAGTGTGTTGCACTG 1290
Oy 1151 ATGTGCGCGTGTGCTGCTCTCTGCTGAGCGCCAAAGATGAGGAAAGTGTGTTGCACTG 1210
Db 1291 ATGTGCGCGTGTGCTGCTCTCTGCTGAGCGCCAAAGATGAGGAAAGTGTGTTGCACTG 1350
Oy 1211 TGTGTCAGCGTGTGCTGCTCTCTGCTGAGCGCCAAAGATGAGGAAAGTGTGTTGCACTG 1270
Db 1351 TGTGTCAGCGTGTGCTGCTCTCTGCTGAGCGCCAAAGATGAGGAAAGTGTGTTGCACTG 1410
Oy 1271 AAGAGCGCTTATGACCGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1330
Db 1411 AAGAGCGCTTATGACCGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1470
Oy 1331 TCTCTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1390
Db 1471 TCTCTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1530
Oy 1391 GGAGAGCGTGAAGAGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1450
Db 1531 GGAGAGCGTGAAGAGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1590
Oy 1451 GTGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1510
Db 1591 GTGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1650
Oy 1511 ACCCGCAGCGCAGAACCAATCAAAATGAAGTGAATTCCTTCCCA 1553
Db 1651 ACCCGCAGCGCAGAACCAATCAAAATGAAGTGAATTCCTTCCCA 1693

RESULT 7
AX224379 1600 bp DNA linear PAT 10-SEP-2001
LOCUS AX224379
DEFINITION Sequence 1 from Patent WO0160990.
ACCESSION AX224379

VERSION AX224379.1 GI:15554631
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Rastelli, L.
TITLE Novel sphingosine kinases
JOURNAL Patent: WO 0160990-A 1 23-AUG-2001;
Curagen Corporation (US); GENENTECH, INC. (US)
FEATURES
source location/Qualifiers
BASE COUNT 265 a 492 c 531 g 311 t 1 others
ORIGIN
Query Match 96.1%; Score 1511.6; DB 6; Length 1600;
Best Local Similarity 98.7%; Pred. No. 1.4e-271;
Matches 1534; conservative 0; Mismatches 19; Indels 1; Gaps 1;
Oy 1 GCGCCACAGCGCGCGCTGCGAGCGCGCGCTGCGGCGACGACCGGATAGAGAGTGAAGGAGG 60
Db 46 GCGCCACAGCGCGCGCTGCGAGCGCGCGCTGCGGCGACGACCGGATAGAGAGTGAAGGAGG 105
Oy 61 AGCG 119
Db 106 AGCG 165
Oy 120 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 179
Db 166 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 225
Oy 180 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 239
Db 226 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 285
Oy 240 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 299
Db 286 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 345
Oy 300 GAACACAGCG 359
Db 346 GAACACAGCG 405
Oy 360 CATGCTGAGAGCG 419
Db 406 CATGCTGAGAGCG 465
Oy 420 GGAAGCG 479
Db 466 GGAAGCG 525
Oy 480 AGCTTCTCTTGAACCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
Db 526 AGCTTCTCTTGAACCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 585
Oy 540 CTGACAGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
Db 586 CTGACAGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
Oy 600 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
Db 646 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 705
Oy 660 GGAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 719
Db 706 GGAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 765
Oy 720 CTTGCTGCTGAGCG 779
Db 766 CTTGCTGCTGAGCG 825

D	551	TGAGGCTGAATTCCTTCACGCTGATGTCACCTGAGCGGGGAACCAACGCGGGGAGCT	610
O	318	GGTACGGTGGGAGAGCTGGGGCCCTGGGAGCGCTGGGAGTCATGTCGTGGAGAGGGCT	377
D	611	GGTGGGTGGAGAGAGCTGGGGCCCTGGAGACCTCTGCTGTCATGTCGTGGAGAGGGCT	670
O	378	GATGCAAGAGCTGGGAGAACGGGCTCATGAGCGGCTGACTGGGAGACCCGATCCAGAA	437
D	671	GATGCAAGAGCTGGTGAACGGGCTCATGAGCGGCTGACTGGGAGACCCGATCCAGAA	730
O	438	GCCCCGTGTAGCTTCCACGACGAGCTCTGGCAACGCGCTGGGACGCTCTTGAACCATTA	497
D	731	GCCCCGTGTAGCTTCCACGACGAGCTCTGGCAACGCGCTGGGACGCTCTTGAACCATTA	790
O	498	TGCTGGCTATGACAGAGTCACCAATGAAAGACCTCTGACCAACTCAGACGCTATTGCTGTG	557
D	791	TGCTGGCTATGACAGAGTCACCAATGAAAGACCTCTGACCAACTCAGACGCTATTGCTGTG	850
O	558	CCGCGCGCTGCTGTACACCAATGAACCTGCTGTCTGCAACGAGGGTCTGGGGGCTGCCCT	617
D	851	CCGCGCGCTGCTGTACACCAATGAACCTGCTGTCTGCAACGAGGGTCTGGGGGCTGCCCT	910
O	618	CTTCTCTGTGTCACGCTGGGCTGGGCTTCATTGCTGATGTGACCTAGAGATGAGAA	677
D	911	CTTCTCTGTGTCACGCTGGGCTGGGCTTCATTGCTGATGTGACCTAGAGATGAGAA	970
O	678	GTAACGCGCTGAGGAGAGTGGCTTCACCTCTGAGGACCTTCGCGCTGGGAGCGCT	737
D	971	GTAACGCGCTGAGGAGAGTGGCTTCACCTCTGAGGACCTTCGCGCTGGGAGCGCT	1030
O	738	GCGACACTACCGCGGCGACACTGGGCTACCTCCCTGTAAGGAAAGTGGGTTCCAAAGACAC	797
D	1031	GCGACACTACCGCGGCGACACTGGGCTACCTCCCTGTAAGGAAAGTGGGTTCCAAAGACAC	1090
O	798	TGGCTCCCCCGTTTGGTGGTGCACGAGAGGCGCGGTATGATGATGACACTTGTGGCACTGAGAGA	857
D	1091	TGGCTCCCCCGTTTGGTGGTGCACGAGAGGCGCGGTATGATGATGACACTTGTGGCACTGAGAGA	1150
O	858	GCGACAGCGGCTCTCACTGTGAGACAGTGGTGGCCGACAGAGACTTGTGTCTAATCTTGGCACT	917
D	1151	GCGACAGCGGCTCTCACTGTGAGACAGTGGTGGCCGACAGAGACTTGTGTCTAATCTTGGCACT	1210
O	918	GCTGCACTGCAACTGGGCACTGAGATGTTTGGTGTGCAACCCATGGGCGGCTGTGACGCTGG	977
D	1211	GCTGCACTGCAACTGGGCACTGAGATGTTTGGTGTGCAACCCATGGGCGGCTGTGACGCTGG	1270
O	978	GCTGATGCACTGTTCTACAGTGGGCGGAGAGTGTCTGCTGCAATGCTGTCGCCCTCTT	1033
D	1271	GCTGATGCACTGTTCTACAGTGGGCGGAGAGTGTCTGCTGCAATGCTGTCGCCCTCTT	1330
O	1038	CTTGGGCAATGAGAGGCGACGATATGAGAGTTGATATGCCCTCTACTTGGTATATGTCGC	1099
D	1331	CTTGGGCAATGAGAGGCGACGATATGAGAGTTGATATGCCCTCTACTTGGTATATGTCGC	1390
O	1098	CGTGGTCCGCTTCGCTTGGAGAGCCCAAGAGATGGGAAAGGTGTGTTGACGAGTGGGGA	1157
D	1391	CGTGGTCCGCTTCGCTTGGAGAGCCCAAGAGATGGGAAAGGTGTGTTGACGAGTGGGGA	1450
O	1158	ATTGATGATGTTTACCGAGAGCGCGAGAGGCGACAGTGGCAACCCAACTACTCTGGATGGTCAG	1211
D	1451	ATTGATGATGTTTACCGAGAGCGCGAGAGGCGACAGTGGCAACCCAACTACTCTGGATGGTCAG	1510
O	1218	CGGTTGCGAGAGAGCGCGCGCCACACTGTGGAAGCCCGACGATGGCACCGCAGAAAGGCG	1277
D	1511	CGGTTGCGAGAGAGCGCGCGCCACACTGTGGAAGCCCGACGATGGCACCGCAGAAAGGCG	1570
O	1278	CTTATATACCCCTGGGCGCGGCTGTGCTTATGATGTCTACTTGTAGAGACCTCTCTCTCC	1333
D	1571	CTTATATACCCCTGGGCGCGGCTGTGCTTATGATGTCTACTTGTAGAGACCTCTCTCTCC	1630
O	1338	CTTGGGCTGAGAGGCTCTCTCCACAGCTCTGTGGGGGTGAGAGACTCTCTTGAGAGAG	1397
D	1631	CTTGGGCTGAGAGGCTCTCTCCACAGCTCTGTGGGGGTGAGAGACTCTCTTGAGAGAG	1690

[illegible]

Db 1145 CCGCGGGGCGGAGGAGGAGGCTTTCAGAGCTTCTCCGAGTCACTGACGAGCCCTTTTGG 1204
Oy 257 CTGAGGCTGAATCTCTTTCACGCTGATGAGTCTGAGGCGGAGCAACGAGCGGAGAC 316
Db 1205 CTGAGGCTGAATCTCTTTCACGCTGATGAGTCTGAGGCGGAGCAACGAGCGGAGAC 1264
Oy 317 TGGTCCGCTGAGAGAGCTGAGGCGGAGGAGCTTGTGTTGATGTTCTGAGAGAGGCG 376
Db 1265 TGGTCCGCTGAGAGAGCTGAGGCGGAGGAGCTTGTGTTGATGTTCTGAGAGAGGCG 1324
Oy 377 TGATGACAGAGTGTGTAAGGAGGCTGATGAGAGGCGCTGATGAGAGAGGAGAGGAGAG 436
Db 1325 TGATGACAGAGTGTGTAAGGAGGCTGATGAGAGGCGCTGATGAGAGAGGAGAGGAG 1384
Oy 437 AGCCCTGTGATGCTCCAGAGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAG 496
Db 1385 AGCCCTGTGATGCTCCAGAGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAG 1444
Oy 497 ATGCTGCTGATGAGAGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCG 556
Db 1445 ATGCTGCTGATGAGAGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCG 1504
Oy 557 GCGGCGGCTGCTGTCACCCATGAAACCTGCTGCTGAGAGGCGCTGAGAGGCGCTGAG 616
Db 1505 GCGGCGGCTGCTGTCACCCATGAAACCTGCTGCTGAGAGGCGCTGAGAGGCGCTGAG 1564
Oy 617 TCTTCTGCTGCTGAGAGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCG 676
Db 1565 TCTTCTGCTGCTGAGAGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCG 1624
Oy 677 AGTATGCGGCTGAGAGGAGATGCGCTTCACTTGGGAGGAGGAGGAGGAGGAGGAGGAG 736
Db 1625 AGTATGCGGCTGAGAGGAGATGCGCTTCACTTGGGAGGAGGAGGAGGAGGAGGAGGAG 1684
Oy 737 TGGGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 796
Db 1685 TGGGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1744
Oy 797 TGGGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 856
Db 1745 TGGGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1804
Oy 857 AGCCAGTGGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 916
Db 1805 AGCCAGTGGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1864
Oy 917 TGGGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 976
Db 1865 TGGGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1924
Oy 977 GCGTATGCTGATGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1036
Db 1925 GCGTATGCTGATGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1984
Oy 1037 TGGGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1096
Db 1985 TGGGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2044
Oy 1097 GCGTATGCTGATGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1156
Db 2045 GCGTATGCTGATGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2104
Oy 1157 AGTATGCTGATGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1216
Db 2105 AGTATGCTGATGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2164
Oy 1217 GCGGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1276
Db 2165 GCGGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2224
Oy 1277 CCGTATGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1336
Db 2225 CCGTATGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2284

Oy 1337 CCTAGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1396
Db 2285 CCTAGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2344
Oy 1397 GGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1456
Db 2345 GGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2404
Oy 1457 AGCCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1516
Db 2405 AGCCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2464
Oy 1517 ACCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1554
Db 2465 ACCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2502

RESULT 10
HSA245504 1428 bp mRNA linear PRI 22-MAY-2000
LOCUS HSA245504
DEFINITION Homo sapiens mRNA for sphingosine kinase (gene SPK).
ACCESSION AJ245504
VERSION AJ245504.1 GI:8017375
KEYWORDS sphingosine kinase; SPK gene.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1428)
AUTHORS Van Veldhoven, P.P. and Gijssels, S.
JOURNAL Unpublished
TITLE 2 (bases 1 to 1428)
AUTHORS Van Veldhoven, P.P.
DIRECT SUBMISSION
SUBMITTED (16-AUG-1999) Van Veldhoven, P.P., Molecular Cell Biology,
katholieke universiteit Leuven, K.U.Leuven, Campus Gasthuisberg,
Afd. Farmakologie, Herestraat, B-3000 Leuven, BELGIUM
REMARK Revised by author (17-MAY-2000)
FEATURES
source location/Qualifiers
1..1428
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q24"
1..1146
/gene="SPK"
1..1146
/gene="SPK"
/codon_start=1
/product="sphingosine kinase"
/protein_id="CAB92131.1"
/db_xref="GI:8017375"
/translation="MDPAGPRGVARLVLLINPRGKAKALQFRSHVOPILAEAI
SPTLALTERRNIARELVESEELGRWDALVYVNSDDGIMHEVNGLERPMPETAIQKPL
CSLPAGSGKPLAASLNHYAGVEQVNEDELITLTCRLLSPNLLISHTASGRL
PSVLSLAKETADVLDLESERKRLDEMRITLCTPLRLALRYRGLRATLPPRGSK
TPASPVVQOQVDALVLEBPVPSHMTVDEDEVLVLLHLSLSEMPAEPGR
CAAGVHLFVYRAGVSRAMLRLFLAMEGRIMEXECPYLIVYPVGAFLPEKDKGV
FAVDELWSEVVOGVHPNPFYMWVSGCEVPPSPMKPQMPPEPEPL"
BASE COUNT 247 a 422 c 462 g 297 t
ORIGIN
Query Match 88.1%; Score 1385.6; DB 9; Length 1428;
Best Local Similarity 99.7%; Pred. No. 4.1e-248;
Matches 1388; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 175 CCGGCTGCTGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 234
Db 37 CCGGCTGCTGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 96
Oy 235 AGTCAAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 294
Db 235 AGTCAAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 294


```

/db_xref="GI:9280106"
/translation="MDPAGPRGVLPSCRHVLVILNDRGKGLQLFERSHVDPLAE
AEISFTLMEERHRLARELRSBELGMDLAVYSGMLHVEVNGMLMERDDETATQ
KPLCSLPAGSGNLAASLAINVAGYEYDNTDLTNCRLICRLISPMNLSLTASG
LRLEFSVLSIAMEPLADYDLSEKIRLIGEMRFTLTGLRLAALTLTGRLAYLEVGA
GSRTPSVVYVQGVDAHLVPLEEPVSPVTVVVDLFLVLALLSHLSSEFPAAP
MGRAAAGVMLFVYRAGVSRAMLRLFLAMEKRMHEYECPYLVIISGWDIG"
BASE COUNT      355 a      639 c      650 g      371 t
ORIGIN

Query Match      83.1%; Score 1306.6; DB 9; Length 2015;
Best Local Similarity 92.3%; Pred. No. 2e-233;
Matches 1452; Conservative 0; Mismatches 64; Indels 57; Gaps 5;

OY      1  GCGCCACAGCGCGCGCTCGACGCGCGCGCTGGGCGACCGCTAAGAGAGCTGAAGCAGG 60
DB      498  GCGCCACAGCGCGCGCTCGACGCGCGCGCTGGGCGACCGCTAAGAGAGCTGAAGCAGG 557
OY      61  AGCGCGCGCACAGCGGCGACGCGCC-ACAGCGCGCAGGAGACCCCTGGCAGCGGACCGCG 119
DB      558  AGCGCGCGCACAGCGGCGACGCGCCGACAGCGCGCAGGAGACCCCGGCAAGAGACCGAC 617
OY      120  GGTGAGGTATGATCCAGCGGCGGCGCGCGCGCTGCTCCCGCGCGCTGCCGCGT 179
DB      618  GGTGAGGTATGATCCAGCGGCGGCGCGCGCGCTGCTCCCGCGCGCTGCCGCGGT 677
OY      180  GCTGCTGCTGTAACCGCGCGCGCGGCGCAAGGCGCTTGACGCTTCGCGAGTCA 239
DB      678  GCTGCTGCTGTAACCGCGCGCGCGGCGCAAGGCGCGCTTGACGCTTCGCGAGTCA 737
OY      240  CGTGACGCGCGCTTTGGCTGAGAGCTGAATCTCTTCAACGCTGATGCTACTAGACGCGG 299
DB      728  CGTGACGCGCGCTTTGGCTGAGAGCTGAATCTCTTCAACGCTGATGCTACTAGACGCGG 797
OY      300  GAGCAGCGCGCGGAGCTGTGCGCTCGAGAGAGCTGGGCGCGCTGGAGCGCTCTGCTGT 359
DB      798  GAGCAGCTGCGGCGGAGCTGTGCGCTCGAGAGAGCTGGGCGCGCTGGAGCGCTCTGCTGT 857
OY      360  CATGCTGTGAGAGCGGCGTATGACAGAGTGGTGAAGGCGCTCATGAGCGCGCTGACGTG 419
DB      858  CATGCTGTGAGAGCGAGCTATGACAGAGTGGTGAAGGCGCTCATGAGCGCGCTGACGTG 917
OY      420  GGAGACCGCGCATCCAGAAAGCCCTGTGTAGCTTCCACAGAGCGCTTGGCAACGCGCTGC 479
DB      918  GGAGACCGCGCATCCAGAAAGCCCTGTGTAGCTTCCACAGAGCGCTTGGCAACGCGCTGC 977
OY      480  AGCTTCTCTTGAACCTTATGCTGCTATGAGCAGGTCAACCAATGAAGACTCTCTGACCA 539
DB      978  AGCTTCTCTTGAACCTTATGCTGCTATGAGCAGGTCAACCAATGAAGACTCTCTGACCA 1037
OY      540  CTGACGCGTATTGCTGTGCGCGCGCGCTGCTGTACCCATGAACCTGCTGTCTGCACAC 599
DB      1038  CTGACGCGCGTATTGCTGTGCGCGCGCGCTGCTGTACCCATGAACCTGCTGTCTGCACAC 1097
OY      600  GCGCTTGGGCGCGCTGCGCTTCTCTGTGCTCAGCGCTGGCGCTTGCATTTGCTGATGT 659
DB      1098  GCGCTTGGGCGCGCGCTGCGCTTCTCTGTGCTAAGCGCTGGCGCTTGCATTTGCTGATGT 1157
OY      660  GGACCTAAGAGAGTGAAGATGCGCGCTGTGGGAGATGCGCTTCACTGTGGGACCTT 719
DB      1158  GGACCTAAGAGAGTGAAGATGCGCGCTGTGGGAGATGCGCTTCACTGTGGGACCTT 1217
OY      720  CCGTGGTGGGACCGCGGCGACCTACCGCGCGCGGCGACTGCGCTACCTCCCTGTAGGAAG 779
DB      1218  CCGTGGTGGGACCGCGGCGACTTACCGGCGCGGCGGCGCTTACCTCCCTGTAGGAAG 1277
OY      780  AGTGCGTTCACAGACACTTGCCTCCCGGTGTGTGCTCCAGCAGCGGCGCGGTATGACACA 839
DB      1278  AGCGCGGCTCTAAGACACTGTGTTCCCGGTGTGTGCTCCAGCAGCGGCGCGGTATGACACA 1337
OY      840  CCTTGTGCGCACTGAGGAGCGGACAGTCCCTCTCAGTGAAGAGTGTGCCCGACGAGACTT 899
DB      1338  CCTTGTGCGCACTGAGGAGCGGACAGTCCCTCTCAGTGAAGAGTGTGCCCGACGAGACTT 1397

```

```

OY      900  TGTGCTAGTCTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
DB      1398  GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1457
OY      960  GGGCGCGCTGTGAGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
DB      1458  GGGCGCGCTGTGAGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1517
OY      1020  CATGCTGCTGCGCGCTTCTTCTGCGCATGAGAGAGGCGAGGCTATGAGATGATGCTGCC 1079
DB      1518  CATGCTGCTGCGCGCTTCTTCTGCGCATGAGAGAGGCGAGGCTATGAGATGATGATGCTGCC 1577
OY      1080  CTACTTGTATATGTGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1139
DB      1578  TTACTTGTATATCAT----- 1592
OY      1140  GTTTCAGTGTGATGAGGGAATGATGTTAGCGAGGCGCTGCAAGGCGCAGGTGACACCAAA 1199
DB      1593  -----CAGTGTGATGAGGGAATGATGTTAGCGAGGCGCTGCAAGGCGCAGGTGACACCAAA 1647
OY      1200  CTACTTGTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259
DB      1648  CTACTTGTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1707
OY      1260  GCCACGCGCAGAGAGCCCTTATGACCCCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1319
DB      1708  GCCACCTCCAGAGAGCCCTTATGACCCCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1767
OY      1320  AGGACCTTCTCTCTTCCCTAGGCGTGCAGGCGCTGTCACAGCTCTGTGTGGGGGTGAG 1379
DB      1768  AGGACCTTCTCTCTTCCCTAGGCGTGCAGGCGCTGTCACAGCTCTGTGTGGGGGTGAG 1826
OY      1380  GAGACTCTCTTGGAGAGGAGGTGAGAGGTGAGAGCTATGCTTTGGGGGACAGGCCAGAA 1439
DB      1827  GAGACTCTCTTGGAGAGGAGGTGAGAGGTGAGAGCTATGCTTTGGGGGACAGGCCAGAA 1886
OY      1440  TGAAGTCTT-GGCTGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1498
DB      1887  TGAAGTCTTGGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1942
OY      1499  GTTGTCTGAGAGCCCGCCACCCAGCAACCAATCAATGAAGTACATTTCCAGCGCTG 1558
DB      1943  ATTGTCTGAGAGCCCGCCACCCAGCAACCAATCAATGAAGTACATTTCCAGCAAA 2002
OY      1559  AAAAAAAAAA 1571
DB      2003  AAAAAAAAAA 2015

RESULT 12
AF200328      1173 bp      mRNA      linear      PRI 25-AUG-2000
LOCUS      AF200328
DEFINITION      Homo sapiens sphingosine kinase (SPHK) mRNA, complete cds.
ACCESSION      AF200328
VERSION      AF200328.1
KEYWORDS      GI:9909360
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens.
REFERENCE      1 (bases 1 to 1173)
AUTHORS      Pitson,S.M., D'Andrea,R.J., Vandeleur,L., Morelli,P.A., Xia,P.,
              Gamble,J.R., Vadas,W.A. and Wattenberg,B.W.
              Human sphingosine kinase: purification, molecular cloning and
              characterization of the native and recombinant enzymes
              Biochem. J. 350 Pt 2, 429-441 (2000)
TITLE      JOURNAL MEDLINE PUBMED
              20407120 10947957
REFERENCE      2 (bases 1 to 1173)
AUTHORS      Pitson,S.M., D'Andrea,R.J., Vandeleur,L., Morelli,P.A.B., Xia,P.,
              Gamble,J.R., Vadas,W.A. and Wattenberg,B.W.
TITLE      Direct Submission

```

JOURNAL		Submitted (29-Oct-1999) Human Immunology.IWVS, Frome Road, Adelaide, SA 5000, Australia	
FEATURES		Location/Qualifiers	
source		1..1173	
gene		/organism="Homo sapiens"	
CDS		/db_xref="taxon:9606"	
		/cell_type="vascular endothelial"	
		1..1173	
		/gene="SPHK"	
		19..1173	
		/gene="SPHK"	
		/codon_start=1	
		/product="sphingosine kinase"	
		/protein_id="AA01980.1"	
		/db_xref="GI:99039361"	
		/translation="MDPARGRVLPRPCRVLVLPNPRGGKNAQLFRSHVQPLLAERISPLMLTERNNHARLVAISELGRMALVNSGDLEHNEVNNILMERPMETAIQLKPLSPSPAGSNAALASLNINAGYEQVTEDEDTLCTLLRLRLSPMKLISLTHASGLRLEFSLAMGFIADLVLESEKYLRLGMRMTLCTLRALARTYRGSLATLVGRVRSKSTPSPVVOOGFVPAHLVLEEPFVSHMTVVDDEFLVLLVSHSGSEMPAARMGSCAAVNHLEFVYRAGYSRAMLRLFLAMEGRIMEECPYLVYVPAVFLFEKDDKGMFAVDGLMAVSEVOQVHPNYFMVYGVGEVPEPPSWKPDQMPPEEPL"	
BASE COUNT	188 a .349 c 390 g 246 t		
ORIGIN			
Query Match	74.1%: Score 1165; Db 9; Length 1173;		
Best Local Similarity	99.6%: Pred. No. 5e-207;		
Matches 1168:	Conservative 0; Mismatches 5; Indels 0; Gaps 0;		
Qy	112 GACCGCGGGGTGAGGTATGATCATCCAGCGGGCGCCCGGGGGCTGTCTCCCGGGCC	171	
Db	1 GACCGCGGGGTGAGGTATGATCATCCAGCGGGCGCCCGGGGGCTGTCTCCCGGGCC	60	
Qy	172 TGGCGGTGCTGCTGCTGCTGAACCCGGCGGGCGCAAGGGACTTGTGACTCTTC	231	
Db	61 TGGCGGTGCTGCTGCTGCTGAACCCGGCGGGCGCAAGGGACTTGTGACTCTTC	120	
Qy	232 CGGAGTCAGTGCACGCCCTTTTGGCTGAGGCTGAATATCTCTTACAGCTGATGCTCACT	291	
Db	121 CGGAGTCAGTGCACGCCCTTTTGGCTGAGGCTGAATATCTCTTACAGCTGATGCTCACT	180	
Qy	292 GACCGCGGAAACCAACGCGCGGAGCTGTGCGGTGGAAGAGCTGGCGCGCTGGAGAGCT	351	
Db	181 GACCGCGGAAACCAACGCGCGGAGCTGTGCGGTGGAAGAGCTGGCGCGCTGGAGAGCT	240	
Qy	352 CTGTGTGTCATGTTGTGAGAGCGGGGCTGATGACAGAGTGTGTGAACGGGCTCATGAGACGG	411	
Db	241 CTGTGTGTCATGTTGTGAGAGCGGGGCTGATGACAGAGTGTGTGAACGGGCTCATGAGACGG	300	
Qy	412 CCGAGTGGAGACCGGCATTCAGAAACCCCGTGTAGCTCCACAGCGCTGTGGCAAC	471	
Db	301 CCGAGTGGAGACCGGCATTCAGAAACCCCGTGTAGCTCCACAGCGCTGTGGCAAC	360	
Qy	472 GCGCTGGCAGCTCTCTTGAACCTATATGCTGGATATAGAGAGTGCACCAATGAGACCTC	531	
Db	361 GCGCTGGCAGCTCTCTTGAACCTATATGCTGGATATAGAGAGTGCACCAATGAGACCTC	420	
Qy	532 CTGACCAACTGCACGCTATATGCTGTGCGCGCGGCTGCTGTACACCCATGAACTGCTGTCT	591	
Db	421 CTGACCAACTGCACGCTATATGCTGTGCGCGCGGCTGCTGTACACCCATGAACTGCTGTCT	480	
Qy	592 CTGACACAGGCTTCGGGGCTGCGCCTCTTCTCTGTGTGTGCACGCTGGCGCTTGATT	651	
Db	481 CTGACACAGGCTTCGGGGCTGCGCCTCTTCTCTGTGTGTGCACGCTGGCGCTTGATT	540	
Qy	652 GGTGATGTGGACCTAGAGATGAGAAATATGGGCTGTGGGGAGAGATGGCTTCACTGT	711	
Db	541 GGTGATGTGGACCTAGAGATGAGAAATATGGGCTGTGGGGAGAGATGGCTTCACTGT	600	
Qy	712 GGGACCTTCCTGAGCTTGGCAGCGCTGGCAACCTAACGGCGGCAATGGGCTACTCCCT	771	
Db	601 GGGACCTTCCTGAGCTTGGCAGCGCTGGCAACCTAACGGCGGCAATGGGCTACTCCCT	660	

QY	772	GTATGAAGAGTGGGTTCTCAAGACACCTGCTCCTCCGTTGTGGTCTCACAGAGGCCCGGTA	831
Db	661	GTATGAAGAGTGGGTTCTCAAGACACCTGCTCCTCCGTTGTGGTCTCACAGAGGCCCGGTA	720
QY	832	GATGCAACACTTGTGGCCACTGGAGAGGCACTGGCCCTTCACACTGAGAGAGTGGTGGCCGAC	891
Db	721	GATGCAACACTTGTGGCCACTGGAGAGGCACTGGCCCTTCACACTGAGAGAGTGGTGGCCGAC	780
QY	892	GAGGACTTGTGTCTAGTCCCTGCGACTGCTGCGACTGCGACCTGGGCAAGTGAATTTTGTCT	951
Db	781	GAGGACTTGTGTCTAGTCCCTGCGACTGCTGCGACTGCGACCTGGGCAAGTGAATTTTGTCT	840
QY	952	GCACCCATGGCCCGCTTGTGCAGTGGCCCTCATGCAATCTGTTTACGCTGGCGGGCGAGTGG	1011
Db	841	GCACCCATGGCCCGCTTGTGCAGTGGCCCTCATGCAATCTGTTTACGCTGGCGGGCGAGTGG	900
QY	1012	TCTGTGCGCATGCTGCTGCGCCCTTCTCCGTGACATGAGAGAGGAGCATATATGAGATAT	1071
Db	901	TCTGTGCGCATGCTGCTGCGCCCTTCTCCGTGACATGAGAGAGGAGCATATATGAGATAT	960
QY	1072	GAAATGCCCTCTCTGTATATATGTGCCCGCTGTGCTCCCTTCCGCTTGGAGCCCAAGATGAG	1131
Db	961	GAAATGCCCTCTCTGTATATATGTGCCCGCTGTGCTCCCTTCCGCTTGGAGCCCAAGATGAG	1020
QY	1132	AAAGGTGTGTTTGCAGTGGATGGGAAATTTGATGGTTAGCGAGAGCCGCTGACAGGCCAGGTG	1191
Db	1021	AAAGGTGTGTTTGCAGTGGATGGGAAATTTGATGGTTAGCGAGAGCCGCTGACAGGCCAGGTG	1080
QY	1192	CACCCAAACTCTCTGTGATGATGACAGCGGTTGCGCTGGAGCCCGCCGCGACGTGGAAACCC	1251
Db	1081	CACCCAAACTCTCTGTGATGATGACAGCGGTTGCGCTGGAGCCCGCCGCGACGTGGAAACCC	1140
QY	1252	CAGCAGATGGCCACCGCCAGAGAGCCGCTTATGA	1284
Db	1141	CAGCAGATGGCCACCGCCAGAGAGCCGCTTATGA	1173
RESULT 13			
LOCUS	AX127642	1155 bp	DNA
DEFINITION	Sequence 2 from Patent W00131029.		
ACCESSION	AX127642		
VERSION	AX127642.1	GI:14134308	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
TITLE	Allen, J., Gosluk, M., Melendez, A.J. and Takacs, L.		
JOURNAL	Human sphingosine kinase gene		
FEATURES	Patent: WO 0131029-A 2 03-MAY-2001;		
source	WARNER-LAMBERT COMPANY (US)		
	Location/Qualifiers		
	1..1155		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	185 a 349 c 382 g 239 t		
ORIGIN			
Query Match	73.4%; Score 1155; DB 6; Length 1155;		
Best Local Similarity	100.0%; Pred. No. 3.6e-205;		
Matches 1155:	Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	130	ATGAGTTCAGAGGGGGGCGCCCGGGGGCGTGTCCCGGGCGCTGGCGGGCTGCTGTGCTG	189
Db	1	ATGAGTTCAGAGGGGGGCGCCCGGGGGCGTGTCCCGGGCGCTGGCGGGCTGCTGTGCTG	60
QY	190	CTGAACCCCGCGCGCGCAAGGCAAGCCCTTGAGCTCTTCCGAGTACAGCTCACGCC	249
Db	61	CTGAACCCCGCGCGCGCAAGGCAAGCCCTTGAGCTCTTCCGAGTACAGCTCACGCC	120
QY	250	CTTTTGGCTGAGGCTGAATCTCTTTCACGCTGATGCTACTAGAGCGCGCAACACGCC	309


```
|||||
Db 121 CTTTGGGTGAGGCTGAATCTCTTCAAGCTGATGCTACAGCGGCGGAGAACCGCG 180
Oy CGGAGCTGAGGCTGAGGAGAGCTGGGCGCTGGGAGCGCTGGTGCTCATGTGCTGA 369
Db 181 CGGAGCTGAGGCTGAGGAGAGCTGGGCGCTGGGAGCGCTGGTGCTCATGTGCTGA 240
Oy 370 GACGGGCTGATGACAGAGGTGAGTGAACGGGCTCATGAGACCGCTGACAGGAGCGCC 429
Db 241 GACGGGCTGATGACAGAGGTGAGTGAACGGGCTCATGAGACCGCTGACAGGAGCGCC 300
Oy 430 ATCCAGAAAGCCCTGTGAGCTCCAGAGGCTGTGGCAAGCGCTGGACGCTTCTTG 489
Db 301 ATCCAGAAAGCCCTGTGAGCTCCAGAGGCTGTGGCAAGCGCTGGACGCTTCTTG 360
Oy 490 AACCATTTATGCTGAGTATGAGAGGCTACACATGAAAGACCTCTTACCACTGACGCTA 549
Db 361 AACCATTTATGCTGAGTATGAGAGGCTACACATGAAAGACCTCTTACCACTGACGCTA 420
Oy 550 TTGCTGTCCCGCGGCTGCTGTCACCCATGAACCTGCTGCTGCTGACACGCGCTTCGGGG 609
Db 421 TTGCTGTCCCGCGGCTGCTGTCACCCATGAACCTGCTGCTGCTGACACGCGCTTCGGGG 480
Oy 610 CTGCGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 669
Db 481 CTGCGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Oy 670 AGTGAGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729
Db 541 AGTGAGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Oy 730 GCAGGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 789
Db 601 GCAGGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Oy 790 AAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 849
Db 661 AAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Oy 850 CTGAGGAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909
Db 721 CTGAGGAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Oy 910 CTGAGGAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
Db 781 CTGAGGAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Oy 970 GCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1029
Db 841 GCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Oy 1030 CGCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1089
Db 901 CGCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Oy 1090 TATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1149
Db 961 TATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Oy 1150 GATGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1209
Db 1021 GATGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Oy 1210 ATGTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1269
Db 1081 ATGTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Oy 1270 GAAGAGCCCTTATGA 1284
Db 1141 GAAGAGCCCTTATGA 1155
```

RESULT 14

```
BC004112
LOCUS BC004112 1192 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, similar to sphingosine kinase 1, clone IMAGE:3832587,
mRNA, partial cds.
ACCESSION BC004112
VERSION BC004112.1 GI:13278662
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 1192).
AUTHORS Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLES Strausberg, R.
JOURNAL Direct Submission
Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: gcgaps@email.nih.gov
Email: gcgaps@email.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbio.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kellerman and Anuradha Madan
FEATURES
source
location/Qualifiers
1..1192
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3832587"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH-MGC_20"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
<1..873
/codon_start=1
/product="Similar to sphingosine kinase 1"
/protein_id="AAH04112.1"
/db_xref="GI:13278662"
/translation="GTRETAIOLKPLCSIPAGSGNALASLNHYAGVEQVFNEDLLTNC
TLILCRRLSPMNLISHTASGLRFSVLSIAMGPIDAVDLESEKRYRLGPMRLTIGT
FLRLAALPTVGRGLAYLPVGRVSKTPASPVYNOGPDALHLPLEPVSHPMYVPD
EDFVLVLAHLISHIGSEMFAPMPGCAAGVNHLPYRAGVSRAMLRLPLAMEGRHM
EYEDPYLVVPAVAFRLPEPKDKGVFAVDGLMSEAVOGVHPYFMVSGVPEPP
SWKPDQMPPEPPL"
CDS
BASE COUNT 245 a 345 c 357 g 245 t
ORIGIN
Query Match 73.3%; Score 1153.4; DB 9; Length 1192;
Best Local Similarity 99.9%; Pred. No. 7.2e-205;
Matches 1154; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 419 GGGAGAGCGCCATCCAGAGGCGCTGTGAGCTCCACAGGCTGTGGCAAGCGCGTGG 478
Db 8 GGGAGAGCGCCATCCAGAGGCGCTGTGAGCTCCACAGGCTGTGGCAAGCGCGTGG 67
Oy 479 CAGCTTCCTTGAACCATTTATGCTGCTATGAGCAGGTCCACATGAAGACTCTTGACCA 538
Db 68 CAGCTTCCTTGAACCATTTATGCTGCTATGAGCAGGTCCACATGAAGACTCTTGACCA 127
Oy 539 ACTGACAGCTATTGCTGTGCGCGCGCTGCTGTACACATGAGAACTGCTGTGACACA 598
Db 128 ACTGACAGCTATTGCTGTGCGCGCGCTGCTGTACACATGAGAACTGCTGTGACACA 187
Oy 599 CGGCTTCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 658
```



```
* 2878 4283: contig of 1406 bp in length
* 4284 4383: gap of 100 bp
* 4384 5447: contig of 1064 bp in length
* 5448 5547: gap of 100 bp
* 5548 6776: contig of 1229 bp in length
* 6777 6876: gap of 100 bp
* 6877 8215: contig of 1339 bp in length
* 8216 8315: gap of 100 bp
* 8316 9961: contig of 1646 bp in length
* 9962 10061: gap of 100 bp
* 10062 12224: contig of 2163 bp in length
* 12225 12324: gap of 100 bp
* 12325 13794: contig of 1470 bp in length
* 13795 13894: gap of 100 bp
* 13895 16857: contig of 2963 bp in length
* 16858 16957: gap of 100 bp
* 16958 18266: contig of 1309 bp in length
* 18267 18366: gap of 100 bp
* 18367 20190: contig of 1824 bp in length
* 20191 20290: gap of 100 bp
* 20291 21460: contig of 1170 bp in length
* 21461 21560: gap of 100 bp
* 21561 23407: contig of 1847 bp in length
* 23408 23507: gap of 100 bp
* 23508 25455: contig of 1948 bp in length
* 25456 25555: gap of 100 bp
* 25556 27331: contig of 1776 bp in length
* 27332 27431: gap of 100 bp
* 27432 30406: contig of 2875 bp in length
* 30407 30506: gap of 100 bp
* 30507 32805: contig of 2299 bp in length
* 32806 32905: gap of 100 bp
* 32906 35659: contig of 2754 bp in length
* 35660 35759: gap of 100 bp
* 35760 37974: contig of 2215 bp in length
* 37975 38074: gap of 100 bp
* 38075 40521: contig of 2447 bp in length
* 40522 40621: gap of 100 bp
* 40622 43219: contig of 2598 bp in length
* 43220 43319: gap of 100 bp
* 43320 46161: contig of 2842 bp in length
* 46162 46261: gap of 100 bp
* 46262 48463: contig of 2202 bp in length
* 48464 48563: gap of 100 bp
* 48564 51275: contig of 2712 bp in length
* 51276 51375: gap of 100 bp
* 51376 53522: contig of 2147 bp in length
* 53523 53622: gap of 100 bp
* 53623 56955: contig of 3333 bp in length
* 56956 57055: gap of 100 bp
* 57056 59555: contig of 2500 bp in length
* 59556 59655: gap of 100 bp
* 59656 62713: contig of 3058 bp in length
* 62714 62813: gap of 100 bp
* 62814 66519: contig of 3706 bp in length
* 66520 66619: gap of 100 bp
* 66620 68930: contig of 2311 bp in length
* 68931 69030: gap of 100 bp
* 69031 72724: contig of 3694 bp in length
* 72725 72824: gap of 100 bp
* 72825 76092: contig of 3268 bp in length
* 76093 76192: gap of 100 bp
* 76193 80143: contig of 3951 bp in length
* 80144 80243: gap of 100 bp
* 80244 84957: contig of 4714 bp in length
* 84958 85057: gap of 100 bp
* 85058 88424: contig of 3367 bp in length
* 88425 88524: gap of 100 bp
* 88525 93428: contig of 4904 bp in length
* 93429 93528: gap of 100 bp
* 93529 97177: contig of 3649 bp in length
* 97178 97277: gap of 100 bp
* 97278 102713: contig of 5436 bp in length
```

```
* 102714 102813: gap of 100 bp
* 102814 109141: contig of 6328 bp in length
* 109142 109241: gap of 100 bp
* 109242 114513: contig of 5272 bp in length
* 114514 114613: gap of 100 bp
* 114614 124525: contig of 9912 bp in length
* 124526 124625: gap of 100 bp
* 124626 132364: contig of 7739 bp in length
* 132365 132464: gap of 100 bp
* 132465 139488: contig of 7024 bp in length
* 139489 139588: gap of 100 bp
* 139589 148734: contig of 9146 bp in length
* 148735 148834: gap of 100 bp
* 148835 159465: contig of 10631 bp in length
* 159466 159565: gap of 100 bp
* 159566 172685: contig of 13120 bp in length
* 172686 172785: gap of 100 bp
* 172786 183443: contig of 10658 bp in length.
* 183443
* 183443
location/Qualifiers
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="CTD-2531H7"
/clone_lib="CITD Human BAC"
1. 1059
misc_feature
/note="assembly_fragment"
1160. 2777
misc_feature
/note="assembly_fragment"
2878. 4283
misc_feature
/note="assembly_fragment"
4384. 5447
misc_feature
/note="assembly_fragment"
5548. 6776
misc_feature
/note="assembly_fragment"
6877. 8215
misc_feature
/note="assembly_fragment"
8316. 9961
misc_feature
/note="assembly_fragment"
10062. 12224
misc_feature
/note="assembly_fragment"
12325. 13794
misc_feature
/note="assembly_fragment"
13895. 16857
misc_feature
/note="assembly_fragment"
16958. 18266
misc_feature
/note="assembly_fragment"
18367. 20190
misc_feature
/note="assembly_fragment"
20291. 21460
Query Match 66.9%; Score 1052.8; DB 2; Length 183443;
Best Local Similarity 99.8%; Pred. No. 2e-186;
Matches 1054; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 503 GCTATGACGAGTGCACCAATGAAGACCTCTGACACGACCTGATTCGTGCGGCC 562
|||||
Db 134217 GCTATGACGAGTGCACCAATGAAGACCTCTGACACGACCTGATTCGTGCGGCC 134158
QY 563 GGCTGCTGCACCAATGAAGACCTCTGACACGACCTGCGGCGCTGCGCTTCT 622
|||||
Db 134157 GGCTGCTGCACCAATGAAGACCTCTGACACGACCTGCGGCGCTGCGCTTCT 134098
QY 623 CTGTGCTGCACCTGCGGCGCTTCTGATTCGATTCGACCTAGACCTGAGAGTATC 682
|||||
Db 134097 CTGTGCTGCACCTGCGGCGCTTCTGATTCGATTCGACCTAGACCTGAGAGTATC 134038
QY 683 GCGCTGCGGAGATGCGCTTCTGACCTGCGGACCTCTGCGCTGCGGACCGCTGCGCA 742
|||||
Db 134037 GCGCTGCGGAGATGCGCTTCTGACCTGCGGACCTCTGCGCTGCGGACCGCTGCGCA 133978
QY 743 CTTACCGCGGCGACTGGCTTCTGCTGATGAGAGATGGGTTCCAGACACTGCGCT 802
```

Mon May 19 10:50:37 2003

us-09-937-060a-19.rge

Page 18

|||||
Db 133977 CTTACCGCGGCGAGCTGGCCCTACCTCCCTGTGAGAAAGTGGGTTCCAGAGACACTGCTCCT 133918
OY 803 CCCCCGTTGTGTCCAGAGAGGCCCCGGTAGATGACACACTTGTGCACACTGGAGAGCCAG 862
Db 133917 CCCCCGTTGTGTGTCCAGAGAGGCCCCGGTAGATGACACACTTGTGCACACTGGAGAGCCAG 133858
OY 863 TCGCCCTCTCAGTGAAGAGTGGTGGCCGAGAGAGACTTGTGTGCTAGTCTGGCACTGGCTG 922
Db 133857 TCGCCCTCTCAGTGAAGAGTGGTGGCCGAGAGAGACTTGTGTGCTAGTCTGGCACTGGCTG 133798
OY 923 ACTCGACCTGGGCGAGTGAATGTTTGTCTGACACCCATGGGCGGCTGTGACAGTGGCGTCA 982
Db 133797 ACTCGACCTGGGCGAGTGAATGTTTGTCTGACACCCATGGGCGGCTGTGACAGTGGCGTCA 133738
OY 983 TGCATCTGTTTACGTGGGGGGGAGTGTCTGTGCCATGCTGTGGCGCTCTTCCTG 1042
Db 133737 TGCATCTGTTTACGTGGGGGGGAGTGTCTGTGCCATGCTGTGGCGCTCTTCCTG 133678
OY 1043 CCATGGAGAAAGGCGAGGATATGAGTATGATGCCCCCTACTGTGTATATGTGCCCTG 1102
Db 133677 CCATGGAGAAAGGCGAGGATATGAGTATGATGCCCCCTACTGTGTATATGTGCCCTG 133618
OY 1103 TCGCCTTCCTGTTGAGGCCCAAGATGGAAAGTGTGTTGCAAGTGAAGGGAATTGA 1162
Db 133617 TCGCCTTCCTGTTGAGGCCCAAGATGGAAAGTGTGTTGCAAGTGAAGGGAATTGA 133558
OY 1163 TGGTTAGCAGAGCCCTGTGAGGCGCAGGTGACCCCAACTACTTGTGATGTCAGCGGTT 1222
Db 133557 TGGTTAGCAGAGCCCTGTGAGGCGCAGGTGACCCCAACTACTTGTGATGTCAGCGGTT 133498
OY 1223 GCGTGAAGCCCGCGCCAGCTGGAAGCCCGAGAGATGCCACCGCAGAAAGGCCCTTAT 1282
Db 133497 GCGTGAAGCCCGCGCCAGCTGGAAGCCCGAGAGATGCCACCGCAGAAAGGCCCTTAT 133438
OY 1283 GACCCCTGGGCGCGCTGTGCTTGTGTACTTGTGACAGACCTTCTCTCTTCCCTAGG 1342
Db 133437 GACCCCTGGGCGCGCTGTGCTTGTGTACTTGTGACAGACCTTCTCTCTTCCCTAGG 133378
OY 1343 GCTGCAGGGCCTGTGTACAGCTCCTGTGGGGGTGAGAGACTGCTGTGAGAAAGGTGA 1402
Db 133377 GCTGCAGGGCCTGTGTACAGCTCCTGTGGGGGTGAGAGACTGCTGTGAGAAAGGTGA 133318
OY 1403 GAAGGTGAGGCTATGCTTGTGGGGAGACAGCCAGATGAAGTCTGTGGTCAAGGCCA 1462
Db 133317 GAAGGTGAGGCTATGCTTGTGGGGAGACAGCCAGATGAAGTCTGTGGTCAAGGCCA 133258
OY 1463 GCTGCTGGGGCCAGCTGCTATGTAAGGCTTCTGTGTGAGAGCCGCCACCCA 1522
Db 133257 GCTGCTGGGGCCAGCTGCTATGTAAGGCTTCTGTGTGAGAGCCGCCACCCA 133198
OY 1523 CGAACCAAAATCAATAAAGTGACATTGCCAGCCTG 1558
Db 133197 CGAACCAAAATCAATAAAGTGACATTGCCAGCCTG 133162

Search completed: May 17, 2003, 13:39:06
Job time : 4577 secs

Mon May 19 10:50:38 2003

us-09-937-060a-19.rni

Page 2

```

APPLICANT: Fuente, Juan L.
APPLICANT: Liarena, Francisco J.
APPLICANT: Liras, Paloma
TITLE OF INVENTION: DNA ENCODING CEPHAMYCIN BIOSYNTHESIS
TITLE OF INVENTION: LATE GENES
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: John W. Mallen III
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04801
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mallen III, John W.
REGISTRATION NUMBER: 5,403
REFERENCE/DOCKET NUMBER: 19179
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3905
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 900 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
PCT-US95-04801-3

```

Query Match	3.0%	Score 47.8;	DB 5;	Length 900;
Best Local Similarity	51.1%;	Pred. No. 0.023;		
Matches 112; Conservative	0;	Mismatches 107;	Indels 0;	Gaps 0

Oy	139	GGGGGGGGCCCCGGGGGCGTCGTCCCGGGGCGCTGCGCGCGTCTGTGTGTCTGTGAACCCG	196
Db	353	GGGGCGGACGACCCGGGGCCCGCCCGACGTCGCCACAGAGGTGTACAGTGGACGACGACCCG	412
Oy	199	GGGGGGGCGAAGGGCAAGGCGCTTGTGACTCTTTCGCGAGTCACGTGACGCCCCCTTTGGCT	256
Db	413	GGCAGCGCGCGGGCAGGGCGGCGGACGACCTTGAGCCCCGGCGGGGATTCGCGCGCTTC	472
Oy	259	GAGCGTGAATCTCTCTTCACGCTGTATGCTACTGACGGCGGAGAACGACGCGGGAGGCTG	318
Db	473	CAGATCATCAACGCTGTGGCGGGCGGCTGTCTGTGACCGGGTGGCCAACTTCTCCGGTGGCGCTG	532
Oy	319	GTGCGGTGGAGGAGCTGGGCGCGGTGGAGACGCTCTTGATG	357
Db	533	TGCGACTACCGCGTGTGGACCTGTGCGCGGACGTGGTG	571

RESULT 3
US-08-209-747-1
Sequence 1, Application US/08209747
Patent No. 5733771
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Colgin, Mark
TITLE OF INVENTION: cDNAs Encoding Minor Ampullate Spider
TITLE OF INVENTION: Silk Proteins
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church

```

STATE: Virginia
COUNTRY: USA
ZIP: 22040-3487
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,747
FILING DATE: 14-MAR-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8050
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEetical: NO
ORIGINAL SOURCE:
ORGANISM: Nephila clavipes
TISSUE TYPE: minor ampullate gland
FEATURE:
NAME/KEY: CDS
LOCATION: 183..2675
OTHER INFORMATION: /product= "N. clavipes mino
OTHER INFORMATION: ampullate silk protein"
IS-08-209-747-1

```

Query Match	3.0%;	Score 47.2;	DB 1;	Length 2793;
Best Local Similarity	47.1%;	Pred. No. 0.044;		
Matches 145; Conservative	0;	Mismatches 163;	Indels 0;	Gaps 0;

QY	105	GCACGGGAGAGCCCGGGGTCAGGTTTATGATCTACAGGGGGCCCCCGGGCGGTGCTGCC	164
Db	1331	GCAGGAGCCCGCTGGGGGTGAGAGCTACGGTGGTCAAGTGGGTTACGGTGCC	1390
QY	165	GGGGCCCTGGCGGCTGGTGGTGGTGTGTAACCGGGCGCGCGGCAAGGCCAAGGCTTCGA	224
Db	1391	GGTGGAGAGCTGGTGGCGGCTGCTGCTGTGAGAGCAGAGCTGGAGCGCGCTGGTGTAC	1450
QY	225	GCCTCTTCGGGAGTCAAGCGGAGGCCCTTTTGGCGTGAAGCTGTAAATCTCTTCAAGCTAT	284
Db	1451	GGTAAAGTGGTGGTGTGCTGACGCTGGAGCTGCTGTGAAGCGCAGAGAGCTGGAGGCTACGCT	1510
QY	285	GCTCACTACAGCGCGGGAACACAGCGGGGAGCTGGTGGCTGGAGAGAGCTGGCGCCCTG	344
Db	1511	GGTCAAGTGGATACGGTGGCGGAGAGAGAGCTGGTGGTGTGACGCTGTGCACACAGGA	1570
QY	345	GCACCGCTTGGTGTCTATGCTCTCTGTGAGACGGGCTGATGTCACACAGGTGGTGACGGCTCAT	404
Db	1571	GGCGGAGCGGCTGGTGGTTACGGTAAAGTGGTGGTGGTGAACCTGGTCCCTGCTGGG	1630
QY	405	GGAGCGGC	412
Db	1631	GCAGGTGC	1638

RESULT 4
US-08-458-298-1
Sequence 1, Application US/08458298
Patent No. 5756677
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Colgin, Mark

Mon May 19 10:50:38 2003

us-09-937-060a-19.rni

Page 6

```

GENERAL INFORMATION:
APPLICANT: Pollock, Thomas J
APPLICANT: Mikolajczak, Marcia
APPLICANT: Yamazaki, Motohide
APPLICANT: Thorne, Linda
APPLICANT: Ameritroul, Richard W
TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
FILE REFERENCE: seq list for appl filed from pro. appl
CURRENT APPLICATION NUMBER: us/03/056,867
EARLIER FILING DATE: 1998-06-11
EARLIER APPLICATION NUMBER: 60/049,428
EARLIER FILING DATE: 1997-06-12
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO. 2
LENGTH: 28804
TYPE: DNA
ORGANISM: Sphingomonas sp. S68
US-03-056,867-2

```

Query Match Similarity	2.7%	Score 42.8	DB 3	Length 28804
Best Local Similarity	49.5%	Pred. No. 0.88		
Matches 110	Conservative	0	Mismatches 112	Indels 0
			Gaps	0
QY	164	CGCGGCGCCGCGCGCGTCTGCTGCTCTGAAACCCGCGCGCGCGCGCAAGGCGCTTGC	223	
Db	28003	CGAGCGGCTTGGGGCGGCATCTGTGGCGACCGCGTGGGCTCTACCGGCATGATCTTCA	28062	
QY	224	AGCTCTTCGGAGTACAGTGCACGCCCTTTTGGCTGAGAGCTGAATCTCTTCACGCTGA	283	
Db	28063	CGTTGGCGGCGGTATCTCTCTCCCTTGGCGAGCGCCAGACCTTCCAGTTACCGTGC	28122	
QY	284	TGCTCAGTGCAGCGCGGAGACACAGCGCGGAGAGTGTGCGGTGGAGAGAGCTGGGCGCT	343	
Db	28123	CCATCTTTCGCCACGCTCTCTGGGCGCGCTGATCTCTTCTGGCGAGACCGACCGCGCATCGCT	28182	
QY	344	GGGAGCGCTCTGTGTGTCATGTCTGGAGAGCGGCGTGATGACG	385	
Db	28183	GGGCGCGCATGATCTGCGCTTCTGGGCGTGTCTTATCTGCG	28224	

```

RESULT 11
US-08-804-227C-1
: Sequence 1, Application US/08084227C
: Patent No. 5876991
:
GENERAL INFORMATION:
:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kunstoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46295
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8221

```

```

1 TELECOMMUNICATION INFORMATION:
2 TELEPHONE: 317-276-2459
3 INFORMATION FOR SEQ ID NO: 1:
4 SEQUENCE CHARACTERISTICS:
5 LENGTH: 43380 base pairs
6 TYPE: nucleic acid
7 STRANDEDNESS: single
8 TOPOLOGY: linear
9 MOLECULE TYPE: DNA (genomic)
10 FEATURE:
11 NAME/KEY: CDS
12 LOCATION: 816..14234
13 FEATURE:
14 NAME/KEY: CDS
15 LOCATION: 14351..15945
16 FEATURE:
17 NAME/KEY: CDS
18 LOCATION: 20010..31199
19 FEATURE:
20 NAME/KEY: CDS
21 LOCATION: 31232..36067
22 FEATURE:
23 NAME/KEY: CDS
24 LOCATION: 36249..41774
25
26 US-08-804-227C-1

```

Query Match	2.7%	Score 42.8	DB 2	Length 43280
Best Local Similarity	48.0%	Pred. No. 0.98		
Matches 122	Conservative	0	Mismatches 132	Indels 0
			Gaps	0
OY 145	GGCCCCGGGGGGGCGTCCCGGGCCCTCGCGCGGTGGTGTGAACCGCGGGCG	204		
Db 12177	GGGGCCCTGTGTATGTCTCCGCTGGCGGGCGGAAAGTGAAGCTGTCTCGCGACGAC	12236		
OY 205	GGCAAGGGCAAGGGCTTCGACGCTTCGCGAGTACAGCGCCCTTTGGCTGAGGCT	264		
Db 12237	GGCGCGAAGCGTGCCTACCGCCGCTCAAGTGGACGGCGGGCTCTGCCCGCCCGCG	12296		
OY 265	GAAATCTCTTCACGCTGATGTCATGTAGGGGGCGAAACAGCGCGGAGTGTGTGG	324		
Db 12297	CGCGCGTCTGTGGGGGTGTCTGGCGCCCGAACCGGAGCGCGTCTCCCGCGACGCGCGCTG	12356		
OY 325	TGCGAAGAGAGCGCGCGCGCTGAGACGCTGTGTGTGCTATCTGTGAAACGGGTATAC	384		
Db 12357	GGGGATCAACCGGACGGGTGGGCCCAACCGGCTGTACCCGAGGTGGCGGCGCTGGCAAG	12416		
OY 385	GAGGTGTGTGAACGG	398		
Db 12417	GGCGTGTGGCGCGG	12430		

RESULT 12
US-08-949-155-49/C
Sequence 49, Application US/08949155
Patent No. 6271436
GENERAL INFORMATION:
APPLICANT: Piedrahita, Jorge A
APPLICANT: Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DUREE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 13:39:12 ; Search time 246 Seconds
(without alignments)
8248.263 Million cell updates/sec

Title: US-09-937-060A-19

Perfect score: 1573

Sequence: 1 gccacacgcgcgcctcgcg.....gcttgaataaaaaaaaaa 1573

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 810007 seqs, 64469091 residues

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCIT_NEM_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCIT_NEM_PUB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1511.6	96.1	1600	US-09-784-810A-1	Sequence 1, Appli
2	1155	73.4	1155	US-09-970-516-1	Sequence 1, Appli
3	853.2	54.2	1759	US-09-784-810A-3	Sequence 3, Appli
4	768	48.8	1149	US-09-970-516-5	Sequence 5, Appli
5	387.6	24.6	394	US-09-954-456-1756	Sequence 156, App
6	291	18.5	480	US-09-783-590-9248	Sequence 9248, Ap
7	262.8	16.7	1857	US-09-970-516-3	Sequence 3, Appli
8	262.8	15.7	2380	US-09-817-676A-13	Sequence 13, Appli
9	246.8	15.7	2698	US-09-817-676A-11	Sequence 11, Appli
10	197.4	12.5	199	US-09-796-692-2905	Sequence 2905, Ap
11	197.4	12.5	199	US-10-040-862-2805	Sequence 2905, Ap
12	193	12.3	296	US-10-015-219-658	Sequence 658, App
13	193	12.3	296	US-09-777-564-658	Sequence 658, App
14	44.8	2.8	88421	US-09-976-059-1	Sequence 1, Appli
15	43.8	2.8	2898	US-09-737-149-3	Sequence 3, Appli
16	43.8	2.8	2916	US-09-737-149-5	Sequence 5, Appli
17	43.8	2.8	3132	US-09-737-149-7	Sequence 7, Appli
18	43.8	2.8	4372	US-09-993-811-1	Sequence 1, Appli
19	43.8	2.8	6823	US-09-989-920-16	Sequence 16, Appli

20	43.6	2.8	594	US-10-123-155-10	Sequence 10, Appli
21	42.8	2.7	2862	US-09-888-615-7	Sequence 7, Appli
22	42.6	2.7	471	US-09-918-995-11975	Sequence 11975, A
23	42	2.7	1131	US-09-934-778-1	Sequence 1, Appli
24	42	2.7	4403	US-09-880-107-2402	Sequence 2402, Ap
25	41.6	2.6	520	US-10-184-644-332	Sequence 332, App
26	41.6	2.6	520	US-10-184-644-332	Sequence 332, App
27	41.4	2.6	1874	US-09-323-998D-33	Sequence 33, Appli
28	41.2	2.6	1903	US-09-832-830A-621	Sequence 621, App
29	41.2	2.6	1990	US-09-765-205-39	Sequence 39, Appli
30	40.8	2.6	629	US-09-968-433-54	Sequence 54, Appli
31	40.4	2.6	2436	US-10-063-547-99	Sequence 99, Appli
32	40.4	2.6	2436	US-10-174-530-309	Sequence 309, App
33	40.4	2.6	2436	US-10-176-758-309	Sequence 309, App
34	40.4	2.6	2436	US-10-063-616-99	Sequence 99, Appli
35	40.4	2.6	2436	US-10-175-737-309	Sequence 307, App
36	40.4	2.6	2436	US-10-063-502-99	Sequence 99, Appli
37	40.4	2.6	2436	US-10-173-706-309	Sequence 309, App
38	40.4	2.6	2436	US-10-175-738-309	Sequence 309, App
39	40.4	2.6	2436	US-10-175-752-309	Sequence 309, App
40	40.4	2.6	2436	US-10-176-482-309	Sequence 309, App
41	40.4	2.6	2436	US-10-176-757-309	Sequence 309, App
42	40.4	2.6	2436	US-10-176-913-309	Sequence 309, App
43	40.4	2.6	2436	US-10-180-552-309	Sequence 309, App
44	40.4	2.6	2436	US-10-180-557-309	Sequence 309, App
45	40.4	2.6	2436	US-10-173-700-309	Sequence 309, App

ALIGNMENTS

RESULT 1
US-09-784-810A-1
Sequence 1, Application US/09784810A
Patent No. US20020082203A1
GENERAL INFORMATION:
APPLICANT: RASTELLI, LUCA
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
TITLE OF INVENTION: SAME
FILE REFERENCE: 10716-08
CURRENT APPLICATION NUMBER: US/09/784,810A
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/182,360
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 60/191,261
PRIOR FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1600
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (1)
OTHER INFORMATION: a, t, c, g, other or unknown
US-09-784-810A-1
Query Match
Best Local Similarity 96.1%; Score 1511.6; DB 10; Length 1600;
Matches 1534; Conservative 0; Mismatches 19; Indels 1; Gaps 1;
1 GCCCCACACGGCCCTGGACGCGCCCTGGACGCGCGGATAGGAGCTGAAGCAGG 60
|||||
46 GCCCCACACGGCCCTGGACGCGCCCTGGACGCGCGGATAGGAGCTGAAGCAGG 105
|||||
61 AGCG 119
|||||
106 AGCG 165
|||||
120 GGTGCGAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 179
|||||
166 GGTGCGAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 225
|||||

OY 180 GCTGCTGCTGCTGAAACCCGGGGGGGGAAGGCAAGGCTTGCAGCTTCCTGGAGTCA 239
DB 226 GCTGCTGCTGCTGAAACCCGGGGGGGGAAGGCAAGGCTTGCAGCTTCCTGGAGTCA 285
OY 240 CTTGACACCCCTTTTGGCTGAGGCTGAATTCCTTCACGCTGATGCTCAGAGCGG 299
DB 286 CTTGACACCCCTTTTGGCTGAGGCTGAATTCCTTCACGCTGATGCTCAGAGCGG 345
OY 300 GAACCAAGCCGGGAGCTGCTGCTGCTGAGAGAGCTGGGCGGCTGGAGCGCTTGTG 359
DB 346 GAACCAAGCCGGGAGCTGCTGCTGCTGAGAGAGCTGGGCGGCTGGAGCGCTTGTG 405
OY 360 CATGCTGAGAGAGGCTGATGACAGAGCTGCTGAAACGGGCTCAGAGCGGCTGACTG 419
DB 406 CATGCTGAGAGAGGCTGATGACAGAGCTGCTGAAACGGGCTCAGAGCGGCTGACTG 465
OY 420 GGAGACCGGCTCAGAGAGCCCTGTGTAGCTCCAGAGGCTCTGGCAAGCGCTGGC 479
DB 466 GGAGACCGGCTCAGAGAGCCCTGTGTAGCTCCAGAGGCTCTGGCAAGCGCTGGC 525
OY 480 AGCTTCTGAAACCATTTATGCTGCTATGAGAGAGTCAACCAATGAAGACCTCTGACCA 539
DB 526 AGCTTCTGAAACCATTTATGCTGCTATGAGAGAGTCAACCAATGAAGACCTCTGACCA 585
OY 540 CTGACACCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 599
DB 586 CTGACACCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645
OY 600 GGGCTTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
DB 646 GGGCTTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
OY 660 GGACCTAGAGAGTGAAGATATCGGCTGCTGGGGAGATCGGCTTCACTCTGGGACCTT 719
DB 706 GGACCTAGAGAGTGAAGATATCGGCTGCTGGGGAGATCGGCTTCACTCTGGGACCTT 765
OY 720 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 779
DB 766 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
OY 780 AGTGGGTTTCAAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 839
DB 826 AGTGGGTTTCAAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
OY 840 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 899
DB 886 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
OY 900 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
DB 946 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
OY 960 GGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1019
DB 1006 GGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065
OY 1020 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079
DB 1066 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
OY 1080 CTACTTGGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1139
DB 1126 CTACTTGGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1185
OY 1140 GTTTCAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1199
DB 1186 GTTTCAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1245
OY 1200 CTACTTGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1259
DB 1246 CTACTTGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1305

OY 1260 GCCACCGCCAGAGAGAGCCCTTATGACCCCTGGGCGGCTGTGCTTACTGCTACTGCTC 1319
DB 1306 GCCACCGCCAGAGAGAGCCCTTATGACCCCTGGGCGGCTGTGCTTACTGCTACTGCTC 1365
OY 1320 AGGACCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1379
DB 1366 AGGACCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1425
OY 1380 GAGACTCTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1439
DB 1426 GAGACTCTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1485
OY 1440 TGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1499
DB 1486 TGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1545
OY 1500 TTTGTTCTGAG 1553
DB 1546 TTTGTTCTGAG 1599

RESULT 2
US-09-970-516-1
Sequence 1, Application US/09970516
Patent No. US20020099029A1
GENERAL INFORMATION:
APPLICANT: No. US20020099029A1artis AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-31617
CURRENT APPLICATION NUMBER: US/09/970,516
CURRENT FILING DATE: 2001-10-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1155
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1155)
OTHER INFORMATION:
US-09-970-516-1

Query Match 73.4%; Score 1155; DB 10; Length 1155;
Best Local Similarity 100.0%; Pred. No. 2, 2e-313;
Matches 1155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 130 ATGGATTCAG 189
DB 130 ATGGATTCAG 60
OY 190 CTGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 249
DB 61 CTGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
OY 250 CTTTGGCTGAGGCTGAATTCCTTCACTGATGCTGACGAGAGAGAGAGAGAGAGAGAGAGAG 309
DB 121 CTTTGGCTGAGGCTGAATTCCTTCACTGATGCTGACGAGAGAGAGAGAGAGAGAGAGAGAG 180
OY 310 CCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 369
DB 181 CCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
OY 370 GAGGAGCTGATGACAG 429
DB 241 GAGGAGCTGATGACAG 300
OY 430 ATCCAG 489
DB 301 ATCCAG 360
OY 490 AACCATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 549

```

Db 361 AACCATTTATGCTGCTATGAGCAGGTACACCAATGAAGACCTCTGACCACTGACAGCTA 420
Oy 550 TTGCTGTGCGCGCGGCTGTGTCACCCATGAACCTCTCTCTGACACGGCTTCGGGG 609
Oy 421 TTGCTGTGCGCGCGGCTGTGTCACCCATGAACCTCTCTCTGACACGGCTTCGGGG 480
Oy 610 CTGCGCCCTCTCTCTGCTGCTACGCTGCGCTGGGGCTTCTGATGCTGATGACCTAGAG 669
Db 481 CTGCGCCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Oy 670 AGTGAAGATATGCGCGCTGTGCGGAGATGCGCTTCTACTCTGCGCACTTCTGCTGCTG 729
Db 541 AGTGAAGATATGCGCGCTGTGCGGAGATGCGCTTCTACTCTGCGCACTTCTGCTGCTG 600
Oy 730 CGAGCGCTGCGGACCTACCGCGCGCGGACCTGCGCTACTCTCTCTGATAGAGAGTGGTTC 789
Db 601 CGAGCGCTGCGGACCTACCGCGCGCGGACCTGCGCTACTCTCTCTGATAGAGAGTGGTTC 660
Oy 790 AAGACACCTGCTCCCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 849
Db 661 AAGACACCTGCTCCCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Oy 850 CTGAGAGAGCCAGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 909
Db 721 CTGAGAGAGCCAGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Oy 910 CTGAGAGAGCCAGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 969
Db 781 CTGAGAGAGCCAGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840
Oy 970 CGAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1029
Db 841 CGAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Oy 1030 CGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1089
Db 901 CGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
Oy 1090 TATGTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1149
Db 961 TATGTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Oy 1150 GATGGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1209
Db 1021 GATGGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Oy 1210 ATGCTCAGCGGTTGCGTGAGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1269
Db 1081 ATGCTCAGCGGTTGCGTGAGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Oy 1270 GAAGAGCCCTTATGA 1284
Db 1141 GAAGAGCCCTTATGA 1155

```

RESULT 3 US-09-784-810A-3

```

; Sequence 3, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 3
; LENGTH: 1759
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-784-810A-3

```

```

Query Match          54.2%; Score 853.2; DB 10; Length 1759;
Best local similarity 73.9%; Pred. No. 6.3e-229;
Matches 1158; Conservative 0; Mismatches 388; Indels 22; Gaps 5;

```

```

Oy 3 CCCACAGCCGCGCCCTGCGAGCGCCCGCTGGGAGCAGCCGATGAAGAGCTGAAGCAGAG 62
Db 196 CCCACAGCGGCTGCGAGCGCCCGCTGGGAGCAGCCGATGAAGAGCTGAAGCAGAG 255
Oy 63 CCGCCCGCCAGCGGCGCGCCCGCTGGGAGCAGCCGATGAAGAGCTGAAGCAGAG 118
Db 256 CCGCCCGCTTACTCTGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 315
Oy 119 GGGTCGAGGTTATGATTCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 178
Db 316 GGGACCTGCTATGAGACAGTAAGATGCCCTTCGAGGACTGCTCCACGCGCATGACAG 375
Oy 179 TGCTGTGCTGCTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 238
Db 376 TGCTGTGCTGCTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 435
Oy 239 ACCTGACGCGCCCTTTTGGCTGAGGCTGAATCTCTTACGCTGATGCTCAGTGAACG 298
Db 436 GTGTGAGCGCTTCTCTGAGGAGGAGAGATTAACCTTTAACTGATTAACCGAGAG 495
Oy 299 GGAACACGCGCGCGGAGCTGCTGCGGAGAGCTGCGGCGCGCGCGCGCGCGCGCGCG 358
Db 496 AGAACCATGCGAGGAGAGCTGCTGCTGCGAGAGAGTGGGTACTGCGAGCGCGCTG 555
Oy 359 TCATGTCTGAGAGCGGCTGATGACAGAGGTGTGAACCGCGCTCATGAGAGCGGCTG 418
Db 556 TCATGTCTGAGAGCGGCTGATGATGAGAGGTGTGAACCGCGCTCATGAGAGCGGCTG 615
Oy 419 GGGAGCGCGCGCATTCAGAGCGCGCTGTGAGCTCCACAGCGCTGCTGCAACCGCG 478
Db 616 GGGAGCGCGCATTCAGAGCGCGCTGTGAGCTCCCTGAGAGCGCTGCGCAAGCGCG 675
Oy 479 CAGCTTCTTGAACCATTAATGCTGCTATGAGAGTCAACCAATGAAGACCTTCTGACA 538
Db 676 CAGCTTCTTGAACCATTAATGCTGCTATGAGAGTCAACCAATGAAGACCTTCTGACA 735
Oy 539 ACTGACAGCTATTTGCTGCGCGCGCGCTGCTGACCATGAACGCTGCTCTGACACA 598
Db 736 ACTGACAGCTATTTGCTGCGCGCGCGCTGCTGACCATGAACGCTGCTCTGACACA 795
Oy 599 CGGCTTCTGCGGCGCTGCGCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 658
Db 796 CGGCTTCTGCGGCGCTGCGCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 855
Oy 659 TGGACCTAGAGTGAAGATATCGGCTGTGGGAGATGAGCTTCACTCTGCGGACCT 718
Db 856 TGGACCTAGAGTGAAGATATCGGCTGTGGGAGATGAGCTTCACTCTGCGGACCT 915
Oy 719 TCTGCTGCTGCGGCGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGAGAA 778
Db 916 TCTGCTGCTGCGGCGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGAGAA 975
Oy 779 GAGTGGCTTCAAGACAGCTGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTG 838
Db 976 GAGTGGCTTCAAGACAGCTGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTG 1032
Oy 839 ACCTTGTGCACTGAGAGAGCAGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 898
Db 1033 ACCTTGTGCTGAGAGAGCAGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1092
Oy 899 TTGTGCTAGTCTGCGACTGCTGCGACTGCGACCTGCGGAGTGAAGTGTGCTGACACA 958
Db 1093 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1152

```

Mon May 19 10:50:38 2003

us-09-937-060a-19.rnpb

Page 4

QY	959	TGGGCCCTGTACACTTCAGCCTCATGCTATGTTTACGTGGGGGAGAGTGTCTG	1018
Db	1153	TGGGGCGGTGAGGCTGGTGTATGATCTGTGTACGTACGTGGGGGTGTCAAGG	1212
QY	1019	CAATGTGTGGGCGCTCTCTGGCCATGGAAGAAGCAGCATATGAGATGAATGCC	1078
Db	1213	GTGGCGTGTGGGCGCTCTCTGGCCATGGAAGAAGCAGCATATGAGATGAATGTC	1272
QY	1079	CGTACTGTGTATATGTGCGCCGTGGTCCCTTCCGCTTGGAGCCCAAGAGATGGAAAGTG	1138
Db	1273	CATACCTGTGTATATGTGCGCCGTGGTCCCTTCCGCTTGGAGCCCAAGAGAGAGGCG	1332
QY	1139	TGTTTGCAGTGTGATGGGAAATTGATGTTAGCAGAGGCGGTGACAGGCGCAGGTGACCCAA	1198
Db	1333	TGTTTCTGTGTGATGGAGAGCTGATGTATGTGAAGCTGTGAGGCCAAGTGCACCCAA	1392
QY	1199	ACTACTTGTGATGTGTGACGGTGCCTGGAGCCCGCCGCGCAGCTGGAAGCCGCCACAA	1258
Db	1393	ACTACTTGTGATGTGTGTGGCAGCAGAGATGTCCCATCCGGCGGGACCTCCCGCGCGG	1452
QY	1259	TGGCACCAGGCAAGAAGCCCT---TATGAGCCCTGGCGCGGTCTGTGCTTAATGCTTAC	1315
Db	1453	GCGCAGCTCTCAAGAAAGACATACTCTGTGCTTCTTGTCTACTGTGTAGGCTAGACATGG	1512
QY	1316	TTTGAGGAGCCCTTCTCTCTTCCCTAGGGCTGCAGGGCCGTCCACAGCTCTGTGGGGGT	1375
Db	1513	GACCTTCCCGCCGACCCACCCCTCTGTGATGTGGAGGCTTATTCTTAAAGTTCCATGTGAAGCT	1572
QY	1376	GGAGGAGACTCTCTGTGGAAGAGGCTGAGAAGGTGAGAGCTATGCTTTGGGGGAGACGCC	1435
Db	1573	GGTGGGAGACCCCTGTGAGAGAAAGCT--AGAAGGTGGGGCTATGACTTGGAAAGAAAGGCT	1630
QY	1436	AGATGAATCTCTGTGTAGAGGCCAGCT-----GGCTAGGCCAGCGTGCCTAT	1485
Db	1631	TTTACTTTCAGTTAGTAGTAATCATCCCAAGTAGAGGCGCTGTGGCTGACCAAGTTGATATT	1690
QY	1486	GTAAGCCCTTTAGTTTGTCTGTGAGCCGCCGCCACAGAACCAATTCAAATTAAGTGA	1545
Db	1691	AGAAGACATTTCCCTTTCCTTTTAGGAGACCTTCCCTGGGAACCAATTCAAATTAAGTGA	1750
QY	1546	CATTCCCA	1553
Db	1751	CTTTTCCA	1758

```

RESULT 4
US-09-970-516-5
Sequence 5, Application US/0970516
Patent No. US20020099029A1
GENERAL INFORMATION:
APPLICANT: NO. US20020099029A1ar1s AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-11617
CURRENT APPLICATION NUMBER: US/09/970,516
CURRENT FILING DATE: 2001-10-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 1149
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1149)
OTHER INFORMATION:
US-09-970-516-5

```

Query Match 48.88; Score 766; DB 10; Length 1149;
Best Local Similarity 79.88; Pred. No. 3.9e-205;
Matches 919; Conservative 0; Mismatches 220; Indels 3; Gaps 1

[illegible]

DB 1078 ATGCTGTGAGCAGAGATGCCCATCCGCGGAGCTCCCGGGGGGCGACCTCCA 1137
QY 1270 GAAGACCCCTTA 1281
DB 1138 GAAGAACCATAA 1149

RESULT 5

US-09-954-456-1756/c
; Sequence 1756, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 1756
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1756

Query Match 24.6%; Score 387.6; DB 10; Length 394;
Best Local Similarity 99.0%; Pred. No. 8,3e-99;
Matches 390; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1165 GTTAGCGAGCCGCTGCGAGGCCAGGTGCACCCAACTACTTCTGTGATGTCAGCGGTTGC 1224
DB 394 GTTAGCGAGCCGCTGCGAGGCCAGGTGCACCCAACTACTTCTGTGATGTCAGCGGTTGC 335
QY 1225 GTGAGAGCCCGCCGCGAGTGAAGCCCGACAGATGCCACCGCCAGAAAGGCCCTATGA 1284
DB 334 GTGAGAGCCCGCCGCGAGTGAAGCCCGACAGATGCCACCGCCAGAAAGGCCCTATGA 275
QY 1285 CCCCTGGGCGCGCTGTGCTTAGTGTACTTGCAGAGCCCTTCCCTCCCTAGGCG 1344
DB 274 CCCCTGGGCGCGCTGTGCTTAGTGTACTTGCAGAGCCCTTCCCTCCCTAGGCG 215
QY 1345 TGCAGGCGCTGTGCACAGCTCCTGTGGGGGTGAGAGACTCCTCTGAGAGGGTGAGA 1404
DB 214 TGCAGGCGCTGTGCACAGCTCCTGTGGGGGTGAGAGACTCCTCTGAGAGGGTGAGA 155
QY 1405 AGGTGAGGCTAGCTTTGGGGGACAGGCCAATGAAGTCTGGGTGAGAGGCCAGC 1464
DB 154 AGGTGAGGCTAGCTTTGGGGGACAGGCCAATGAAGTCTGGGTGAGAGGCCAGC 95
QY 1465 TGCGTGGGCGCAGCTCCATATGTAAGGCCCTTCTAGTTTGTGAGAGCCCGCCAGC 1524

DB 94 TGCGTGGGCGCAGCTCCATATGTAAGGCCCTTCTAGTTTGTGAGAGCCCGCCAGC 35
QY 1525 AACCAATCCCAATATAAGTACATTTCCCGAGCCTG 1558
DB 34 AACCAATCCCAATATAAGTACATTTCCCGAGCCTG 1

RESULT 6

US-09-783-590-9248
; Sequence 9248, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haselbine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16,2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 9248
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: (68)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (193)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (220)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (221)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (236)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (289)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (297)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (322)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (337)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (350)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (356)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (367)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (382)

```

OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (384)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (386)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (391)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (400)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (412)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (451)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-9248

```

```

Query Match 18.5%; Score 291; DB 10; Length 480;
Best Local Similarity 86.4%; Pred. No. 9, 2e-72;
Matches 400; Conservative 0; Mismatches 55; Indels 8; Gaps 8;

```

```

OY 266 AATCTCTTCAAGCTGATGCTCACTGAGCGGGAACCAACGCGGGGAGCTGCTGCGGT 325
DB 10 AATCTCTTCAAGCTGATGCTCACTGAGCGGGAACCAACGCGGGGAGCTGCTGCGGT 69
OY 326 CGAGGAGCTGCGGCGGCTGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 385
DB 70 CGAGGAGCTGCGGCGGCTGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129
OY 386 AG-CTGCTGACGCGG-TCATGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442
DB 130 AGAGTGTGACAGCGGCTGATGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 189
OY 443 TGTCTACCTCCAGCAGGCTCTGCGCAACGCGCTGCGCACTTCTTGAACCATATGCTG 502
DB 190 TGTCTACCTCCAGCAGGCTCTGCGCAACGCGCTGCGCACTTCTTGAACCATATGCTG 249
OY 503 GCTATG-AGCAGGTGACCAATG-AAGACCTCTGACCAACCTGACAGCTATGCTGTG-CC 559
DB 250 GTTATGAGAGAGGTGACCAATGAAAGACCTCTGACCAANTGACCTGATGCTGCTGCTG 309
OY 560 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 618
DB 310 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 369
OY 619 TTTCTGCTG-CTCTGAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 677
DB 370 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
OY 678 GTATCGGCGCTGCGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 430 GTATCGGCGCTGCGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 472

```

```

RESULT 7
US-09-970-516-3
; Sequence 3, Application US/09970516
; Patent No. US20020099029A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020099029A1a1a1s AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent version 3.1
; SEQ ID NO 3
; LENGTH: 1857

```

```

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1857)
OTHER INFORMATION:
US-09-970-516-3

```

```

Query Match 16.7%; Score 262.8; DB 10; Length 1857;
Best Local Similarity 63.7%; Pred. No. 9, 4e-64;
Matches 399; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

```

```

OY 147 CCCCCTGGGCGCTGCTCCCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 206
DB 408 CACCCCTGACCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 467
OY 207 CAAGGGCAAGGCTTGCAGCTCTTCCGAGTCACTGACGACCCCTTTTGGCTGAGGCTGA 266
DB 468 TCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 527
OY 267 AATCTCTTCAAGCTGATGCTCACTGAGCGGGAACCAACGCGGGGAGCTGCTGCGGTG 326
DB 528 GCTGCTCTTCAAGCTGATGCTCACTGAGCGGGAACCAACGCGGGGAGCTGCTGCGGTG 587
OY 327 GGAGGAGCTGCGGCGGCTGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 386
DB 588 GCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 647
OY 387 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 446
DB 648 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 707
OY 447 TACCTCTTCAAGCTGATGCTCACTGAGCGGGAACCAACGCGGGGAGCTGCTGCGGTG 506
DB 708 CATCTCTTCAAGCTGATGCTCACTGAGCGGGAACCAACGCGGGGAGCTGCTGCGGTG 767
OY 507 TGAAGAGTCAACCAATGAAAGCTCTGACCAACCTGACAGCTATGCTGTGCTGCTGCTG 566
DB 768 TGAAGAGTCAACCAATGAAAGCTCTGACCAACCTGACAGCTATGCTGTGCTGCTGCTGCTG 827
OY 567 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 626
DB 828 TGCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 887
OY 627 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686
DB 888 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 947
OY 687 TCTGAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 746
DB 948 CTGAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1007
OY 747 CCGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 772
DB 1008 CCGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1033

```

```

RESULT 8
US-09-817-676A-13
; Sequence 13, Application US/09817676A
; Patent No. US2002042101A1
; GENERAL INFORMATION:
; APPLICANT: Spiegell, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian sphingosine kinase type 2 isoforms, cloning,
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent version 3.1
; SEQ ID NO 15
; LENGTH: 1857

```

SEQ ID NO 13
LENGTH: 2380
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (7)..(1860)
PUBLICATION INFORMATION:
TITLE: Molecular cloning and functional characterization of a
JOURNAL: J Biol. Chem.
VOLUME: 275
ISSUE: 26
PAGES: 19513-19520
DATE: 2000-06-30
DATABASE ACCESSION NUMBER: AF245447
DATABASE ENTRY DATE: 2000-06-27
US-09-817-676a-13

Query Match 16.7%; Score 262.8; DB 10; Length 2380;
Best Local Similarity 63.7%; Pred. No. 9.9e-64;
Matches 399; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

OY 147 CCCCCGGGGGCTGCTCCCGCGCCCTGCGGCTGCTGCTGAACCCGCGCGCG 206
DB 414 CACCCCTGACCTGCTACCTCGCGCCCGCGGTTGCTTATGCTCAATCCCTTGGGGG 473
OY 207 CAAGGGCAAGGCTTGCAGCTTCCGGAGTCACCTGACGCCCTTTGGTGGAGCTGA 266
DB 474 TCGGGGCTGCTGCTGCGAGTGATGTAAGAACCCGCTGCTCCAGATCTCTGAAGCTGG 533
OY 267 AATCTCTTCAACGCTGATGCTACTGAGCGGGAACACCGCGCGAGCTGCTCGGTC 326
DB 534 GCTGCTCTCAACCTCATTCAGACAGAACAGAACACCGCGCGAGCTGCTCAAGG 593
OY 327 GAGAGAGCTGGGCGGCTGGAGAGCTCTGCTGCTCATGCTGAGAGCGGCTGATGACGA 386
DB 594 GCTGAGCTGAGTGAAGTGAAGTGCATGCTCAGGCTCTCGGAGAGCGGCTGCTCATGA 653
OY 387 GGTGCTGAACGGGCTCATGAGAGCGGCTGCTGAGAGACCGCATCCAGAACCCCTGTG 446
DB 654 GGTGCTGAACGGGCTCATGAGATCGCCGCTGAGAGAGAGAGTGAAGATCCCTGTGG 713
OY 447 TAGCTTCCAGCAGGCTTGGCAACGCGCTGACCTTCTTGAACATTTATGCTGCTA 506
DB 714 CATCTTCCCTGCGGCTGGGGAAGCGCTGCGCGAGAGAGTGAACACGACGGGGGAT 773
OY 507 TGAGAGCTCAACCAATGAAGCTCTGTAACCACTGACAGCTATTGCTGCGCGGCT 566
DB 774 TGAGCAGACCCCTGGGCTGCGACCTGTTGCTCAACTGCTGCTGCTGCGGGTGG 833
OY 567 GCTGTCACCAATGAACGCTGCTCTGACACAGCGCTTGGGGGCTGCGCTTCTCTCT 626
DB 834 TGGCCACCACTGAGAGCTGCTCTGTAACGCTGCGCTGCGCTGCTGTTTCTCTT 893
OY 627 GCTCAGCTGCGGCTGGGCTTCTATGCTGATGCTGAGCTAGAGAGTGAAGATATGCGG 686
DB 894 CCTGCTCTGCTGGGCTTGGGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 953
OY 687 TCTGGGAGAGATGCGCTTCACTCTGAGGACCTTCTGCTGCTGAGAGCGCTGCGACCTA 746
DB 954 CTGAGGAGAGTGGGCTTCACTGAGGACGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 1013
OY 747 CCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 772
DB 1014 CCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1039

RESULT 9
US-09-817-676a-11
Sequence 11, Application US/09817676A
Patent No. US20020042101A1
GENERAL INFORMATION:

APPLICANT: Spiegel, Sarah
APPLICANT: Kohama, Takahashi
TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
TITLE OF INVENTION: Expression and Methods of Use Thereof
FILE REFERENCE: 00170/HG
CURRENT APPLICATION NUMBER: US/09/817, 676A
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/194, 318
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 2698
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (387)..(2237)
PUBLICATION INFORMATION:
TITLE: Molecular cloning and functional characterization of a
JOURNAL: J Biol. Chem.
VOLUME: 275
ISSUE: 26
PAGES: 19513-19520
DATE: 2000-06-30
DATABASE ACCESSION NUMBER: AF245448
DATABASE ENTRY DATE: 2000-06-27
US-09-817-676a-11

Query Match 15.7%; Score 246.8; DB 10; Length 2698;
Best Local Similarity 62.1%; Pred. No. 3e-59;
Matches 389; Conservative 0; Mismatches 237; Indels 0; Gaps 0;

OY 147 CCCCCGGGGGCTGCTCCCGCGCCCTGCGGCTGCTGCTGCTGAACCCGCGCGCG 206
DB 797 CACCCCTGATGCTGCTGCGCGGAGAGCCGCTGCTCATATGCTCAATCCCTTGGGGG 856
OY 207 CAAGGGCAAGGCTTGCAGCTTCCGGAGTCACCTGACGCCCTTTGGTGGAGCTGA 266
DB 857 GCGGGGCTGCGGCTGCGAGCGCTGATGAGACCGCTGCTGCTGCTGCTGCTGCTGCTGCT 916
OY 267 AATCTCTTCAACGCTGATGCTACTGAGCGGGAACACCGCGCGAGCTGCTCGGTC 326
DB 917 GCTGCTCTCAACCTCATTCAGACAGAACAGAACACCGCGCGAGCTGCTCAAGG 976
OY 327 GAGAGAGCTGGGCGGCTGGAGAGCTCTGCTGCTCATGCTGAGAGCGGCTGATGACGA 386
DB 977 GTTAAAGCTGAGTGAAGTGAAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
OY 387 GGTGCTGAACGGGCTCATGAGAGCGGCTGCTGAGAGACCGCATCCAGAACCCCTGTG 446
DB 1037 GGTGCTGAATGAGGCTCTTGAAGGCGCAGCTGAGAGAGATGCTGCGGATGCCATTTGG 1096
OY 447 TAGCTTCCAGCAGGCTTGGCAACGCGCTGACCTTCTTGAACATTTATGCTGCTA 506
DB 1097 TGTCTTCCCTGCTGAGATGCGGCAATGCGCTAGCTGGGCGGATGAGACCATATGCGGGT 1156
OY 507 TGAGAGCTCAACCAATGAAGCTCTGTAACCACTGACAGCTATTGCTGCTGCGCGGCT 566
DB 1157 TGAGCAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216
OY 567 GCTGTCACCAATGAACGCTGCTCTGACACAGCGCTTGGGGGCTGCGCTTCTCTCT 626
DB 1217 CAGCCATCTCTGAGAGCTGCTCTGTAACGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1276
OY 627 GCTCAGCTGCGGCTGGGCTTCTATGCTGATGCTGAGCTAGAGAGTGAAGATATGCGG 686
DB 1277 CCTGCTCTGCTGGGCTTGGGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 1336
OY 687 TCTGGGAGAGATGCGCTTCACTCTGAGGACCTTCTGCTGCTGAGAGCGCTGCGACCTA 746
DB 1337 CCTGAGGAGAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1396

Mon May 19 10:50:38 2003

us-09-937-060a-19.inpb

Page 8

OY 747 CCGGCGCGGCTGACCTGCTG 772
Db 1397 CCGTGAGCGCTCTCTACCTCCCG 1422

RESULT 10
US-09-796-692-2905

; Sequence 2905, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Galger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077, 001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 2905
; LENGTH: 199
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-2905

Query Match 12.5% Score 197.4; DB 9; Length 199;
Best Local Similarity 99.5% Pred. No. 1.2e-45;
Matches 198; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1043 CCATGAGAGAGGCGCATATGAGTATGATGCCCCCTACTTGGTATATGCCCCGTG 1102
Db 1 CCAATGAGAGAGGCGCATATGAGTATGATGCCCCCTACTTGGTATATGCCCCGTG 60
OY 1103 TCGCCCTCCGCTTGGAGGCCCAAGAGTGTCTTTGCACTGAGTGGGAATTGA 1162
Db 61 TCGCCCTCCGCTTGGAGGCCCAAGAGTGTCTTTGCACTGAGTGGGAATTGA 120
OY 1163 TGGTTAGGAGGCGCTGAGGCGCAGGTGACCAAACTACTTGTGATGTCAGCGGTT 1222
Db 121 TGGTTAGGAGGCGCTGAGGCGCAGGTGACCAAACTACTTGTGATGTCAGCGGTT 180
OY 1223 GCGTGAAGCGCCCGCCGAG 1241
Db 181 GCGTGAAGCGCCCGCCGAG 199

RESULT 11

US-10-040-862-2905
; Sequence 2905, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Galger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 2905
; LENGTH: 199
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-2905

Query Match 12.5% Score 197.4; DB 9; Length 199;
Best Local Similarity 99.5% Pred. No. 1.2e-45;
Matches 198; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1043 CCATGAGAGAGGCGCATATGAGTATGATGCCCCCTACTTGGTATATGCCCCGTG 1102
Db 1 CCAATGAGAGAGGCGCATATGAGTATGATGCCCCCTACTTGGTATATGCCCCGTG 60
OY 1103 TCGCCCTCCGCTTGGAGGCCCAAGAGTGTCTTTGCACTGAGTGGGAATTGA 1162
Db 61 TCGCCCTCCGCTTGGAGGCCCAAGAGTGTCTTTGCACTGAGTGGGAATTGA 120
OY 1163 TGGTTAGGAGGCGCTGAGGCGCAGGTGACCAAACTACTTGTGATGTCAGCGGTT 1222
Db 121 TGGTTAGGAGGCGCTGAGGCGCAGGTGACCAAACTACTTGTGATGTCAGCGGTT 180
OY 1223 GCGTGAAGCGCCCGCCGAG 1241
Db 181 GCGTGAAGCGCCCGCCGAG 199

RESULT 12
US-10-015-219-658
; Sequence 658, Application US/10015219
; Publication No. US2003008299A1


```
NAME/KEY: misc_feature
LOCATION: (39713)..(65800)
OTHER INFORMATION: ORF 14: positive strandedness
NAME/KEY: misc_feature
LOCATION: (65826)..(66530)
OTHER INFORMATION: ORF 15: positive strandedness
NAME/KEY: misc_feature
LOCATION: (66546)..(67370)
OTHER INFORMATION: ORF 16: positive strandedness
NAME/KEY: misc_feature
LOCATION: (67384)..(70059)
OTHER INFORMATION: ORF 17: positive strandedness
NAME/KEY: misc_feature
LOCATION: (70099)..(70662)
OTHER INFORMATION: ORF 18: positive strandedness
NAME/KEY: misc_feature
LOCATION: (70659)..(71906)
OTHER INFORMATION: ORF 19: positive strandedness
NAME/KEY: misc_feature
LOCATION: (73439)..(71964)
OTHER INFORMATION: ORF 20: negative strandedness
NAME/KEY: misc_feature
LOCATION: (74216)..(73563)
OTHER INFORMATION: ORF 21: negative strandedness
NAME/KEY: misc_feature
LOCATION: (75424)..(74213)
OTHER INFORMATION: ORF 22: negative strandedness
NAME/KEY: misc_feature
LOCATION: (75535)..(76464)
OTHER INFORMATION: ORF 23: positive strandedness
NAME/KEY: misc_feature
LOCATION: (78110)..(76449)
OTHER INFORMATION: ORF 24: negative strandedness
NAME/KEY: misc_feature
LOCATION: (79864)..(78107)
OTHER INFORMATION: ORF 25: negative strandedness
NAME/KEY: misc_feature
LOCATION: (81624)..(79861)
OTHER INFORMATION: ORF 26: negative strandedness
NAME/KEY: misc_feature
LOCATION: (81509)..(81682)
OTHER INFORMATION: ORF 27: negative strandedness
NAME/KEY: misc_feature
LOCATION: (82346)..(82062)
OTHER INFORMATION: ORF 28: negative strandedness
NAME/KEY: misc_feature
LOCATION: (82587)..(84446)
OTHER INFORMATION: ORF 29: positive strandedness
NAME/KEY: misc_feature
LOCATION: (84481)..(85548)
OTHER INFORMATION: ORF 30: positive strandedness
NAME/KEY: misc_feature
LOCATION: (85556)..(86945)
OTHER INFORMATION: ORF 31: positive strandedness
NAME/KEY: misc_feature
LOCATION: (85572)..(86803)
OTHER INFORMATION: ORF 32: positive strandedness
NAME/KEY: misc_feature
LOCATION: (87544)..(88420)
OTHER INFORMATION: ORF 33: positive strandedness; N-terminus only
US-09-976-059-1

Query Match 2.8% Score 44.8: DB 9: Length 88421:
Best Local Similarity 45.3% Pred. No. 0.02:
Matches 163: Conservative 0: Mismatches 197: Indels 0: Gaps 0:
```

```
215 AGCGCTTACGCTCTTCGCGAGTCACGTGACGCCCTTTTGGCTGAGCTGAATGCTCT 274
16106 GCGACGTACTGACGCGGGGCTGCTGATCAGCGGAAACCGACCGCGGCTCTCGGC 16047
275 TCAGCTGATGCTCAGTCAGCGCGGACGACCGGAGCTGTGGTGGTGGAGAGC 334
16046 TCGCCCATCAGCTGACGCGGACCGCGGACCGAGATCGAGTGCAGCTCGCCCTGATC 15987
335 TGGCGCGCTGGAGCGCTTGTGTGTCATGTCTGAGAGAGGAGCTGATGACGAGTGTCA 394
15986 TCGACGAACTCGCCGACCTGATACAGCGGCTGTGCGGCGGACGCTGCTGCGCTGCA 15927
395 ACGGCTCATCGACGCGGCTGCTGCTGAGACCGGCTTCAGAGAGCCCTGTGTAGCTTC 454
15926 ACGGCAAGCTGGCGGCGCTGAGCGGCGGCTGACAGGCTGCGGACGAGCTGACCTTC 15867

RESULT 15
US-09-737-149-3
Sequence 3: Application US/09737149
Patent No. US20020077466A1
GENERAL INFORMATION:
APPLICANT: Spademe, Steven K
APPLICANT: Spademe, Kerry E
APPLICANT: Shinkets, Richard A
APPLICANT: Muralidhara, Padigaru
TITLE OR INVENTION: Polypeptides and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-620 CIP
CURRENT APPLICATION NUMBER: US/09/737,149
PRIORITY APPLICATION NUMBER: 60/170,564
PRIORITY FILING DATE: 1999-12-14
PRIORITY APPLICATION NUMBER: 60/173,165
PRIORITY FILING DATE: 1999-12-27
PRIORITY APPLICATION NUMBER: 60/173,362
PRIORITY FILING DATE: 1999-12-27
PRIORITY APPLICATION NUMBER: 60/174,544
PRIORITY FILING DATE: 1999-12-29
PRIORITY APPLICATION NUMBER: 60/174,404
PRIORITY FILING DATE: 2000-01-04
PRIORITY APPLICATION NUMBER: 60/174,962
PRIORITY FILING DATE: 2000-01-07
PRIORITY APPLICATION NUMBER: 60/223,929
PRIORITY FILING DATE: 2000-08-09
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 2898
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2898)
US-09-737-149-3

Query Match 2.8% Score 43.8: DB 10: Length 2898:
Best Local Similarity 47.3% Pred. No. 0.019:
Matches 166: Conservative 0: Mismatches 182: Indels 3: Gaps 1:
```

```

OY 239 ACGTCACGCCCCCTTTGGCTGAGGCTGAATCTCCTTCACGCTGATGCTCAGCTGAGCGGC 298
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 335 AGCTCTGCGAGCTGCACCTTCTGCGGCGGCCACCGAGACCCCGCTGCTCAGCCTGCTGC 394
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 299 GGAACGACGCGGCGGAGTGCTGCGGTGAGAGAGCTGCGCGCTGGGACGCTCTGGTGG 358
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 395 GCGGGAAGCGCGCGCCCTCGGAGCCCCGAAACCATTCACCTGCAGCTGCACTGGG 454
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 359 TCATGCTGAGACGCGCTGATGACAGAGTGTGAGACGCGCTCATGAGC 409
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 455 CCAGCCCCCTGAGAGAGCTGCTGATGTGCTGTGGCGGTGCTGCAGCGGC 505
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: May 17, 2003, 17:20:32
 Job time : 428 secs



GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 09:07:01 : Search time 2408 seconds
(Without alignments)
10579.527 Million cell updates/sec

Title: US-09-937-060A-19

Perfect score: 1573
Sequence: 1 gccacacagccgagccctgcg.....gacctgaaaaaaaaaaaaa 1573

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estlba:*
2: em_estlum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlcv:*
6: em_estlpl:*
7: em_estlro:*
8: em_estl:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1547.8	98.4	2133	11	BC014439 Homo sapi
2	878	55.8	1054	14	BM808698 AGENCOURT
3	845.8	53.8	1137	14	BM810136 AGENCOURT
4	814.6	51.8	892	13	B1860351 AGENCOURT
5	787.8	50.1	888	14	BO647377 AGENCOURT
6	734.8	46.7	785	12	BG680521 AGENCOURT

Result No.	Score	Query Match	Length	DB ID	Description
7	732.8	46.6	772	9	AT769914
8	716.8	45.6	1100	14	BM916526
9	643	40.5	676	10	AW963415
10	636.4	40.5	702	13	BE464487
11	630.6	40.1	875	10	BM006005
12	629	40.0	793	12	BG678689
13	625.4	39.8	662	13	BM386957
14	622	39.5	638	9	AI972156
15	620.2	39.4	809	12	BG280830
16	612.8	39.0	671	10	AV694791
17	606.8	38.6	659	12	AM662445
18	601.8	38.3	658	12	BE891653
19	600	38.1	940	12	BG026631
20	597	38.0	606	14	BQ251614
21	591	37.6	999	14	BO675531
22	582.2	37.0	603	10	BE275818
23	572	36.4	1005	13	BM559257
24	563.2	35.8	852	10	BE274434
25	559	35.5	1075	13	BM557357
26	545.8	34.7	2145	11	BC011432
27	532.6	33.9	842	13	B1648186
28	526.2	33.5	947	13	B1415545
29	510.2	32.4	587	9	AU147968
30	477.8	30.4	720	14	BQ109456
31	476.2	30.3	534	12	BE904632
32	469.2	29.8	856	13	B1757316
33	468	29.8	478	12	BE740866
34	455.6	29.0	1134	14	BM808666
35	449.4	28.6	697	13	B1253500
36	446.4	28.4	1232	13	BM556915
37	441.8	28.1	1115	13	BM550039
38	439	27.9	447	9	AI042283
39	426.2	27.1	914	13	B1823329
40	411.2	26.1	416	13	BM471470
41	407.8	25.9	732	12	BE728564
42	401.4	25.5	733	14	BM934157
43	400.2	25.4	662	12	BE893504
44	397.6	25.3	412	9	AA639414
45	397.2	25.3	618	10	AW916618

ALIGNMENTS

RESULT 1
LOCUS BC014439 2133 bp mRNA linear HTC 19-DEC-2001
DEFINITION Homo sapiens, clone IMAGE:4871343, mRNA.
ACCESSION BC014439
VERSION BC014439.1 GI:17939597
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2133)
TITLES
AUTHORS Strausberg, R.
DIRECT SUBMISSION
JOURNAL Submitted (17-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC/DCRD/PTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Matheson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natsiya van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Zatta.

Clone distribution: MGC clone distribution information can be found through the I.M.G.E. Consortium/URL at: <http://Image.lim.gov>
Series: IRM Plate: 34 Row: 1 Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 11464966
This clone has the following problem: Incomplete processing.

FEATURES

Source

1. 2133
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4871343"
/issue_type="Skin, melanotic melanoma, high MDR."
/clone_lib="NIH-MGC-49"
/lab_host="DH10B-R"
/note="Vector: pOT87"
BASE COUNT 350 a 681 c 719 g 383 t
ORIGIN

Query Match

Best Local Similarity 98.4%; Score 1547.8; DB 11: Length 2133;
Matches 1563; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

OY 1 GGGCCAGAGCGGCGCTGGAGCGCGGCTGGAGCGCGATTAAGACCTTAAGCGCG 60
DB 563 GCGCCAGAGCGGCGCTGGAGCGCGGCTGGAGCGCGATTAAGACCTTAAGCGCG 622
OY 61 AGCG 119
DB 623 AGCG 682
OY 120 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 683 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 742
OY 180 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
DB 743 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 802
OY 240 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 299
DB 803 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 862
OY 300 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 359
DB 863 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 922
OY 360 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 419
DB 923 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 982
OY 420 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 479
DB 983 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1042
OY 480 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 539
DB 1043 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1102
OY 540 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 599
DB 1103 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1162
OY 600 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 659
DB 1163 GGTGAGGTTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1222

OY 660 GGACCTAGAGAGTGAAGATATGCGCGCTGGGAGATGCGCTTACCTGGGCGCTT 719
DB 1223 GGACCTAGAGAGTGAAGATATGCGCGCTGGGAGATGCGCTTACCTGGGCGCTT 1282
OY 720 CTTGCTGCTGGACCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 779
DB 1283 CTTGCTGCTGGACCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1342
OY 780 AGTGCTTCCAGACACCTGCTCCCGCTGGTGTGTCAGACAGGCGCGGTATGAC 839
DB 1343 AGTGCTTCCAGACACCTGCTCCCGCTGGTGTGTCAGACAGGCGCGGTATGAC 1402
OY 840 CTTGCTGCTGGACCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 899
DB 1403 CTTGCTGCTGGACCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1462
OY 900 TGTGCTAGTCTGGACCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 959
DB 1463 TGTGCTAGTCTGGACCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1522
OY 960 GGGCGCGCTGGCG 1019
DB 1523 GGGCGCGCTGGCG 1582
OY 1020 CATGCTGCTGGCG 1079
DB 1583 CATGCTGCTGGCG 1642
OY 1080 CATGCTGCTGGCG 1139
DB 1643 CATGCTGCTGGCG 1702
OY 1140 GTTGGAGTGGGGAATGATGTTAGCGAGCGCGGTGAGCGCGCGCGCGCGCGCGCG 1199
DB 1703 GTTGGAGTGGGGAATGATGTTAGCGAGCGCGGTGAGCGCGCGCGCGCGCGCGCG 1762
OY 1200 CTACTTGTGATGTCACCGGCTGCGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1259
DB 1763 CTACTTGTGATGTCACCGGCTGCGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1822
OY 1260 GCGACCGCGAGAGAGCGCTTATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1319
DB 1823 GCGACCGCGAGAGAGCGCTTATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1882
OY 1320 AGGACCGCTTCTCTTCCCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1379
DB 1883 AGGACCGCTTCTCTTCCCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1942
OY 1380 GAGACTCTCTGAGAGAGCGGTGAGAGCGGTGAGAGCGGTGAGAGCGGTGAGAGCG 1439
DB 1943 GAGACTCTCTGAGAGAGCGGTGAGAGCGGTGAGAGCGGTGAGAGCGGTGAGAGCG 2002
OY 1440 TGAAGTCTGCGGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1499
DB 2003 TGAAGTCTGCGGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2062
OY 1500 TTTGTTCTGAGAGCG 1559
DB 2063 TTTGTTCTGAGAGCG 2122
OY 1560 AAAAAAATAA 1570
DB 2123 AAAAAAATAA 2133

RESULT 2
BM808698 1054 bp mRNA Linear EST 05-MAR-2002
LOCUS BM808698
DEFINITION AGENCOURT.6582622 NIH_MGC-41 Homo sapiens cDNA clone IMAGE:5471258
ACCESSION BM808698
VERSION 5' mRNA sequence.
KEYWORDS BM808698.1 GI:19125509 EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1054)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LCM1979 row: g column: 03
High quality sequence stop: 7
High quality sequence stop: 666.
Location/Qualifiers
FEATURES
source
1..1054
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5471258"
/clone.lib="NIH_MGC_41"
/tissue.type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: This is a
NIH-MGC Library."

BASE COUNT 177 a 317 c 340 g 220 t
ORIGIN
Query Match 55.8%; Score 878; DB 14; Length 1054;
Best Local Similarity 94.0%; Pred. No. 1.6e-173;
Matches 981; Conservative 0; Mismatches 50; Indels 13; Gaps 6;

QY 272 CCTTCAGCGTGTCTCAGTGTGAGCGGCGGAGCGGCGGAGCTGGTCGGGAGG 331
DB 1 CCTTCAGCGTGTCTCAGTGTGAGCGGCGGAGCGGCGGAGCTGGTCGGGAGG 60
QY 332 AGCTGGGCGCTGGAGCGCTGTGTGTCTGTGAGAGCGGCGTGTGACAGAGCTGG 391
DB 61 AGCTGGGCGCTGGAGCGCTGTGTGTCTGTGAGAGCGGCGTGTGACAGAGCTGG 120
QY 392 TGAACGGGCTGTGAGAGCGGCTGTGAGTGTGTGAGAGCGGCGTGTGAGAGG 451
DB 121 TGAACGGGCTGTGAGAGCGGCTGTGAGTGTGTGAGAGCGGCGTGTGAGAGG 180
QY 452 TCCAGAGCGCTGTGAGAGCGGCTGTGAGTGTGTGAGAGCGGCGTGTGAGAGG 511
DB 181 TCCAGAGCGCTGTGAGAGCGGCTGTGAGTGTGTGAGAGCGGCGTGTGAGAGG 240
QY 512 AGGTCAACATGAGAGCGCTGTGAGTGTGTGAGAGCGGCGTGTGAGAGG 571
DB 241 AGGTCAACATGAGAGCGCTGTGAGTGTGTGAGAGCGGCGTGTGAGAGG 300
QY 572 CACCCATGAACCTGTGTCTGTGAGAGCGGCTGTGAGAGCGGCGTGTGAGAGG 631
DB 301 CACCCATGAACCTGTGTCTGTGAGAGCGGCTGTGAGAGCGGCGTGTGAGAGG 360
QY 632 GCGTGGCGTG 691
DB 361 GCGTGGCGTG 420
QY 692 GGGAGATGCGCTTCACTCTGTGGGACCTTCTGTGGTGTGGAGCGGCTTACCGCG 751
DB 751 GGGAGATGCGCTTCACTCTGTGGGACCTTCTGTGGTGTGGAGCGGCTTACCGCG 751

DB 421 GGGAGATGCGCTTCACTCTGTGGGACCTTCTGTGGTGTGGAGCGGCTTACCGCG 480
QY 752 GCGGACTGGCGCTACCTCCCTGTAGGAAGAGTGGTTCAGACACCTGCTCCCGGTTG 811
DB 481 GCGGACTGGCGCTACCTCCCTGTAGGAAGAGTGGTTCAGACACCTGCTCCCGGTTG 540
QY 812 TGGTTCAGAGCGGCGCGGTGTAGATGACACCTTGTGCGACTGTGAGAGCGGCGGTTG 871
DB 541 TGGTTCAGAGCGGCGCGGTGTAGATGACACCTTGTGCGACTGTGAGAGCGGCGGTTG 600
QY 872 ACTGAGACGTGTGTGCGGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTG 931
DB 601 ACTGAGACGTGTGTGCGGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTG 650
QY 932 TGGGCGATGATG 991
DB 661 TGGGCGATGATG 720
QY 992 TCTACGTGC-GGGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1050
DB 721 TCTACGTG 780
QY 1051 AAGGGCAGCATATGAGATGAAATGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1110
DB 781 AAGGGCAGCATATGAGATGAAATGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTG 840
QY 1111 CGCTTGAGCGCCCAAGATGTGGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1168
DB 841 CGCTTGAGCGCCCAAGATGTGGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 900
QY 1169 GC-GAAGCGGCTGTGAGAGCGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1222
DB 901 GCGGAGCGGCGTGTGAGAGCGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 960
QY 1223 GCGTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1278
DB 961 GCGTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
QY 1279 TTATGACCGCTG 1302
DB 1021 TTATGACCGCTG 1044

RESULT 3
BM810136 1137 bp mRNA linear EST 05-MAR-2002
LOCUS
DEFINITION AGENCOURT_6579661 NIH_MGC_41 Homo sapiens cdna IMAGE:547089
5', mRNA sequence.
ACCESSION BM810136
VERSION BM810136.1 GI:19126959
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1137)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LCM1976 row: f column: 10
High quality sequence stop: 623.
Location/Qualifiers
FEATURES
source
1..1137
/organism="Homo sapiens"

Db 181 GGGCTGTCACCCATGAACTGCTGTCTCTGACACGGGCTTGGGGCTTGGCTTCT 240
QY 623 CTGTGCTCAAGCTGGGCTGGGCTTCACTGATGTGACCTAGAGAGTGAAGTATC 682
Db 241 CTGTGCTCAAGCTGGGCTGGGCTTCACTGATGTGACCTAGAGAGTGAAGTATC 300
QY 683 GGGCTGTCAGAGATGCGCTTCACTGATGTGACCTAGAGAGTGAAGTATC 742
Db 301 GGGCTGTCAGAGATGCGCTTCACTGATGTGACCTAGAGAGTGAAGTATC 360
QY 743 CCTACCGCGCGGAGTGGGCTTCACTGATGTGACCTAGAGAGTGAAGTATC 802
Db 361 CCTACCGCGCGGAGTGGGCTTCACTGATGTGACCTAGAGAGTGAAGTATC 420
QY 803 CCCCCCTTGTGTCAGAGAGGCGGCTGATGATGACACCTTGTGACCTAGAGAGC 862
Db 421 CCCCCCTTGTGTCAGAGAGGCGGCTGATGATGACACCTTGTGACCTAGAGAGC 480
QY 863 TGGCTCTGTCAGTGAAGTGGTGGCGGAGAGACTTGTGATGCTGACCTGCTGC 922
Db 481 TGGCTCTGTCAGTGAAGTGGTGGCGGAGAGACTTGTGATGCTGACCTGCTGC 540
QY 923 ACTCGACCTGGGAGATGATGATGCTGTCACCTAGGCGGCTGTCAGCTGCTCA 982
Db 541 ACTCGACCTGGGAGATGATGATGCTGTCACCTAGGCGGCTGTCAGCTGCTCA 599
QY 983 TGCATCTGTTCACTGTCGGGCGGAGTGTCTGTGTCATGCTGCTGCTTCTCTG 1042
Db 600 TGCATCTGTTCACTGTCGGGCGGAGTGTCTGTGTCATGCTGCTGCTTCTCTG 659
QY 1043 CCATGAGAGAGGAGCATATGAGTATGATGATGATGATGATGATGATGATGATG 1102
Db 660 ACATGAGAGAGGAGCATATGAGTATGATGATGATGATGATGATGATGATGATG 719
QY 1103 TCCGCTTCCGCTTGGAGCCCAAGATG-GGAAAGTGTGTTTGCATGATGAGGAATG 1161
Db 720 TCCGCTTCCGCTTGGAGCCCAAGATG-GGAAAGTGTGTTTGCATGATGAGGAATG 779
QY 1162 ATGTTAGGAGAGGCGTGGAGGCGGAGTGGACCAACTTGTGATGATGATGATG 1221
Db 780 ATGTTAGGAGAGGCGTGGAGGCGGAGTGGACCAACTTGTGATGATGATGATG 838
QY 1222 TGGCTGAGAGCGGCGGAGTGGAGAGGCGGAGTGGACCAACTTGTGATGATG 1276
Db 839 TGC-TGGAGAGCGGCGGAGTGGAGAGGCGGAGTGGACCAACTTGTGATGATG 891

RESULT 5
B0647377 888 bp mRNA linear Est 15-JUL-2002
LOCUS AGENCOURT 8414975 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6272071
DEFINITION 5', mRNA sequence.
ACCESSION B0647377
VERSION B0647377.1 GI:21771549
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 888)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLCM2449 row: n column: 08

High quality sequence start: 6
High quality sequence stop: 688.
Location/Qualifiers
1. 888
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6272071"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH-MGC
library."

BASE COUNT 137 a 274 c 284 g 193 t
ORIGIN

Query Match 50.1%; Score 787.8; DB 14; Length 888;
Best Local Similarity 97.1%; Pred. No. 1.2e-154;
Matches 834; Conservative 0; Mismatches 22; Indels 3; Gaps 3;

QY 537 CAACCTGACCTTATGCTGTGCGGCGGCTGTCACCATGACCTGCTGTGCA 596
Db 30 CCACTGACCTTATGCTGTGCGGCGGCTGTCACCATGACCTGCTGTGCA 89
QY 597 CAGGCTTGGGCGGCTGCTTCTCTGTCAGCTTGGGCTTCACTTCTGA 656
Db 90 CAGGCTTGGGCGGCTGCTTCTCTGTCAGCTTGGGCTTCACTTCTGA 149
QY 657 TGTGACCTGAGAGTGAAGTATGCGGCTTGGGAGATGCTTCACTGAGCAC 716
Db 150 TGTGACCTGAGAGTGAAGTATGCGGCTTGGGAGATGCTTCACTGAGCAC 209
QY 717 CTTCGCGCTGTGGAGCGCTTGGAGCGCTACGCGGCGGCTGCTTCTGTAGG 776
Db 210 CTTCGCGCTGTGGAGCGCTTGGAGCGCTACGCGGCGGCTGCTTCTGTAGG 269
QY 777 AAGAGTGGTTCAGAGACACTGCTCCCGTGTGTCACAGAGGCGGCTAGATGC 836
Db 270 AAGAGTGGTTCAGAGACACTGCTCCCGTGTGTCACAGAGGCGGCTAGATGC 329
QY 837 ACACTTGTGCCACTGAGAGAGCCAGTGCCTCTCACTGACAGTGTGCCAGAGA 886
Db 330 ACACTTGTGCCACTGAGAGAGCCAGTGCCTCTCACTGACAGTGTGCCAGAGA 389
QY 897 CTTCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 956
Db 390 CTTCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 449
QY 957 CATGGGCGGCTGTCAGAGCTGCGCTCATGATGCTTCTACGTCGGGCGGAGTGTGC 1016
Db 450 CATGGGCGGCTGTCAGAGCTGCGCTCATGATGCTTCTACGTCGGGCGGAGTGTGC 509
QY 1017 TGGCATGCTGCTGCTTCTTCTGCGCATGAGAGAGGAGCATATGATGATG 1076
Db 510 TGGCATGCTGCTGCTTCTTCTGCGCATGAGAGAGGAGCATATGATGATG 559
QY 1077 CCCCCTACTGATATGTCGCCGCGGCTTCCGCTTGGAGCCCAAGATGGAGAGG 1136
Db 570 CCCCCTACTGATATGTCGCCGCGGCTTCCGCTTGGAGCCCAAGATGGAGAGG 629
QY 1137 TGTGTTGACAGTGAAGGGAATGATGTTAGGAGGCGGTCAGAGGCGGTCACCC 1196
Db 630 TGTGTTGACAGTGAAGGGAATGATGTTAGGAGGCGGTCAGAGGCGGTCACCC 689
QY 1197 AAACCTACTTGTGATGCTGACG-GGTTGCGTGAAGCCCGGCGCAG-CTGGAAGCCCG 1254
Db 690 AAACCTACTTGTGATGCTGACG-GGTTGCGTGAAGCCCGGCGCAG-CTGGAAGCCCG 749

Accession	Sequence	Position
Oy	CAGTGGCCACGGGCGAAGAGAGCCCTTTAGA - CCGCTGGGGCGGGCGTGGACCTTAGTGTCT	1113
Oy	1255 CAGTGGCCACGGGCGAAGAGAGCCCTTTAGA - CCGCTGGGGCGGGCGTGGACCTTAGTGTCT	1113
Dh	750 CAGATGCCACGGCGAAGAGAGCCCTTAGTACCCCGCTGGCGCGCGTGGAGCTCT	809
Oy	1314 ACTGTAGAGACCTCTCCTCCTCCCTAGGGGCGAGGGGCGCTGTCCACAGCTCTGTGGG	1373
Dh	810 ACTGTAGAGACCTCTCCTCCTCCCTAGGGGCGAGGGGCGCTGTCCACAGCTCTGTGGG	869
Oy	1374 GTGGAGAGAGACTCTCTGG 1392	
Dh	870 GGGTGGAAAACACACCTTCG 888	

[illegible]

Accession	Version	Accession	Version
BG680521	BG680521.1	GI:13911918	

Classification: Vertebrata; Euteleostomi;

REFERENCE
1 (bases 1 to 785)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL Unpublished (1999)

COMMENT **Contact:** Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: James Cleaver, M.D.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

FEATURES
Plate: L1AM10612 row: 9 column: 12
High quality sequence stop: 783.
Location/Qualifiers

```
source      1. .785
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
```

```
/clone="IMAGE:4753187"  
/clone_1ib="NCI_CGAP_Skn4"  
/tissue_type="squamous cell
```

```

/lab_host="DH10B (T1 phage-
/note="Organ: skin; Vector:
/site_2: SalI; Cloned unclon

```

Average insert size 1.5kb.
Technologies. Note: this is

ORIGIN	Score
Query Match	45.78: Score 734.8:

Best-Local Similarity 99.18; Pred. No. 1.46
Matches 781; Conservative 0; Mismatches

QY 500 CTGGCTATGAGCAGGTCACCAATGAGACCTCCTGACC
|||||
CTGGCTATGAGCAGGTCACCAATGAGACCTCCTGACC

DD 1 CUGGCAAGACCGATCCAAAGAGAGCCTCATAG

OY 560 GCCGGCTGTGCACCATGAACCTGCTCTTCGCAAG

Db 61 GCCGGCTGTCACCCATGAACTGCTCTCTGCAG

Qy	620	TCTCTGTGCTCAGCCCTGGCCCTGGGCTTCATTGCTGA
Db	121	TCTCTGTGCTCAGCCCTGGCCCTGGGCTTCATTGCTGA

QY 680 ATCGGCTCTGGGGAGATGGCCTTCACTCTGGGCAC

DB 181 ATGGGCGTCTGGGGGAGATGGCGCTTCACTCTCTGGGACCTCTCCCTCGTGTGGCAGCCCTTC 240

OY 740 GCACCTACCGCGGGCGGACCTGGGCTACTCTCCCTGTAGGAAGATGGGTTCCAGACACTTG 799

DB 241 GCACCTACCGCGGGCGGACCTGGGCTACTCTCCCTGTAGGAAGATGGGTTCCAGACACTTG 300

OY 800 CTTCCCGCGTGTGGTTCGACGAGCGCCCGGTAGATGCACACCTTGTCGCACCTGGAGAGAC 859

DB 301 CTTCCCGCGTGTGGTTCGACGAGCGCCCGGTAGATGCACACCTTGTCGCACCTGGAGAGAC 360

OY 860 CAGTGGCCCTCTACCTGAGACAGTGGTGGCGGACGAGACCTTTGTGTAGTCTGGCGACCTG 919

DB 361 CAGTGGCCCTCTACCTGAGACAGTGGTGGCGGACGAGACCTTTGTGTAGTCTGGCGACCTG 420

OY 920 TGCACCTGGACCTGGGCGACGTAGATGTTTCTGTGACCCATGGGCGGCTGTGCACATGGGG 979

DB 421 TGCACCTGGACCTGGGCGACGTAGATGTTTCTGTGACCCATGGGCGGCTGTGCACCTGGG 480

OY 980 TCATGCTATGTTTACGTGCGGGCGGGAGATGTCGTGTCATGCTGCTGCCCTCTTC 1039

DB 481 TCATGCTATGTTTACGTGCGGGCGGGAGATGTCGTGTCATGCTGCTGCCCTCTTC 540

OY 1040 TGGGCGATGGGAGGAGGCGAGCCATATGAGATGAATGACCCTACTTGGTATATGCCCC 1099

DB 541 TGGCGCATGGGAGGAGGCGAGCCATATGAGATGAATGACCCTACTTGGTATATGCCCC 600

OY 1100 TGGTGGCTCTTCGCG -TTGGAGCGCCAAAGATGGGAAAGTGTGTTGGCAGTGGAGGAA 1158

DB 601 TGGTGGCTCTTCGCGCTTTGGAGCGCCAAAGATGGGAAAGTGTGTTGGCAGTGGAGGAA 660

OY 1159 TTGATGTGTATGACGAGCGGTGCGAGGGCCAGGTGACCCAACTACTTCTGGATGTGTAC 1218

DB 661 TTGATGTGTATGACGAGCGGTGCGAGGGCCAGGTGTACC -AAATACTTCTGGATGTGTAC 718

OY 1219 GGTGGCTGTGAGCGCCCGCCGCGACGTGGAAGCCCGACGATGGCACCCCGCAAGAGAGCC 1278

DB 719 GGTGGCTGTGAGCGCCCGCCG -CGAGTGAAGCCCGACGATGGCACCG -CAGAAGAGAGCC 776

OY 1279 TTATGACC 1286

DB 777 TTATGACC 784

RESULT 7

Crack 26-2

AL69914/c			
A1769914	772 bp	mRNA	linear
LOCUS	NCL-GCAP.kid12 Homo sapiens CDNA clone	IMAGE:240431.3	'
DEFINITION	wj30d06.xl	NC_008670.1 NC-CCAP.CDNACGKTNASE	mRNA sequence

SIMILAR TO IR:0888888 0888888 SEFTINOCASINE NAME: 17
 ACCESSION A1769914
 VERSION A1769914.1 GI:5236423

KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens
human,
Homo sapiens

REFERENCE
1 (bases 1 to 772)

AUTHORS: NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE: National Cancer Institute, Cancer Genome Anatomy Project (CCGAP),
Tumor Gene Index

JOURNAL
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: csapbs-r@mail.nih.gov
COMMENT

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution Information can be found at: <http://www.genome.wustl.edu/seqcenter/nci-ccap/>

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrr/image/image.html
Insert length: 1211 Std Error: 0.00

Seq primer: -40UP from Gibco

Crack 26-28-99

REFERENCE
AUTHORS
TITLE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 702)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Db 1 TGAAGACCTCTGACCACTGACACGCTATGCTGTGCCCGGCTGTCTGACCCATGMA 60
Oy 582 CCTGCTGTCTCTGACACAGGCTTGGGGCTGCGCTCTCTCTGTGCTGACGCTGAGC 641
Db 61 CCTGCTGTCTCTGACACAGGCTTGGGGCTGCGCTCTCTCTGTGCTGACGCTGAGC 120
Oy 642 GCGCTTATGCTGATGTGACCTGAGAGAGTATGAGGCTGTGGGGAGATGCG 701
Db 121 GGGCTTATGCTGATGTGAGCTGAGAGAGTATGAGGCTGTGGGGAGATGCG 180
Oy 702 CTTCACCTGTGGGACCTTCTGCTGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 761
Db 181 CTTCACCTGTGGGACCTTCTGCTGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 240
Oy 762 CTTCACCTGTGGGACCTTCTGCTGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 821
Db 241 CTTCACCTGTGGGACCTTCTGCTGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 300
Oy 822 GGGCTTATGCTGATGTGACCTGAGAGAGTATGAGGCTGTGGGGAGATGCG 881
Db 301 GGGCTTATGCTGATGTGAGCTGAGAGAGTATGAGGCTGTGGGGAGATGCG 360
Oy 882 GCGCTTATGCTGATGTGACCTGAGAGAGTATGAGGCTGTGGGGAGATGCG 941
Db 361 GCGCTTATGCTGATGTGAGCTGAGAGAGTATGAGGCTGTGGGGAGATGCG 420
Oy 942 GATGCTTGTGCTGACCCAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 1001
Db 421 GATGCTTGTGCTGACCCAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 480
Oy 1002 GCGGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1061
Db 481 GCGGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Oy 1062 TATGCA-GTATGATGATGATGATGATGATGATGATGATGATGATGATG 1119
Db 541 TATGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Oy 1120 CCAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1175
Db 601 CCAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Oy 1176 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1230
Db 661 GTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Oy 1231 CCGGCTTATGCTGATGTGATGTGATGTGATGTGATGTGATGTGATGTG 1282
Db 721 ACAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Oy 1283 GAGGCTTATGCTGATGTGATGTGATGTGATGTGATGTGATGTGATGTG 1339
Db 781 GAGGCTTATGCTGATGTGATGTGATGTGATGTGATGTGATGTGATGTG 840
Oy 1340 AGGCTTATGCTGATGTGATGTGATGTGATGTGATGTGATGTGATGTG 1374
Db 841 AGGCTTATGCTGATGTGATGTGATGTGATGTGATGTGATGTGATGTG 874

RESULT 12
LOCUS B678689 793 bp mRNA linear EST 01-MAY-2001
DEFINITION 60262448BFL NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749573 5',
mRNA sequence.
ACCESSION B678689
VERSION B678689.1 GI:13910086
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM10602 row: P column: 22
High quality sequence stop: 669.
Location/Qualifiers
1: 793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4749573"
/issue_lib="NCI_CGAP_Skn4"
/issue_type="squamous cell carcinoma"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site: 1; NCI;
Site: 2; Sali; Cloned unidirectionally; primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP library."
BASE COUNT 129 a 244 c 255 g 165 t
ORIGIN
Query Match 40.0%; Score 629; DB 12; Length 793;
Best Local Similarity 93.8%; Pred. No. 1,8e-121;
Matches 744; Conservative 0; Mismatches 35; Indels 14; Gaps 8;
Oy 270 CTCCTTACCTGATGCTGATGATGATGATGATGATGATGATGATGATG 329
Db 1 CTCCTTACCTGATGCTGATGATGATGATGATGATGATGATGATGATG 60
Oy 330 GGAAGTGGGCGGCTGGGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 389
Db 61 GGAAGTGGGCGGCTGGGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
Oy 390 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 449
Db 121 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 180
Oy 450 CTTCCAGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTG 509
Db 181 CTTCCAGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTG 240
Oy 510 GCAAGTACCAATGAAGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTG 569
Db 241 GCAAGTACCAATGAAGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTG 300
Oy 570 GTCACCAATGAAGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTGAGAG 629
Db 301 GTCACCAATGAAGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTGAGAG 360
Oy 570 GTCACCAATGAAGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTGAGAG 629
Db 301 GTCACCAATGAAGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTGAGAG 360
Oy 630 CAGGCTGGGCTGGGCTTCAATGCTGATGTGATGTGATGTGATGTGATGTG 688
Db 361 CAGGCTGGGCTGGGCTTCAATGCTGATGTGATGTGATGTGATGTGATGTG 420
Oy 689 TGGGGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 748
Db 421 TGGGGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 480
Oy 749 GCGGCGAGCTGGGCTTCAATGCTGATGTGATGTGATGTGATGTGATGTG 806
Db 481 GCGGCGAGCTGGGCTTCAATGCTGATGTGATGTGATGTGATGTGATGTG 540
Oy 807 CGTTGTGCTGACAGAGAGCGGCTGATGATGATGATGATGATGATGATGATG 864
Db 541 CGTTGTGCTGACAGAGAGCGGCTGATGATGATGATGATGATGATGATGATG 600
Oy 865 CCTTCTACCTGACAGAGTGTGCGGAGAGAGCT-TTGTGCTAGTCTGTGACAGTGTGCA 923

```

Db      601  CCAATCTACTGACAGTGTGCCCCAGACGACTATCTGTAGTCTGACACTGTGCA 660
Oy      924  CTCGACCTGGG--CAAGGAGATGTTTGTGTGACACCATGGCGCGCTGTG---CACTGGC 978
Db      661  CTCGACCTGGGAGAGTGCAGCTGTACTGTGAGCCCATGGCGGAGTGTGCACCTGGC 720
Oy      979  GTCATGCACTCTGTTTCTAGTGTGCGGCGGAGAGTGTCTGCTG---CCATGCTGTGCGCCTC 1035
Db      721  GTCATGCACTCTGTGTAGTGTGCGGAGGAGATGTATGTCATGCTGCTGTGACGCGCTTC 780
Oy      1036 TTCTGTGCGCATGCG 1048
Db      781  TTCTGTGCGCATGCG 793

RESULT 13
BM386957/c 662 bp mRNA linear EST 17-JAN-2002
LOCUS      BM386957
DEFINITION UI-R-CNI-cj1-c-13-0-UI.s1 UI-R-CNI Rattus norvegicus cDNA clone
ACCESSION  BM386957
VERSION     BM386957.1 GI:18187010
KEYWORDS   EST.
SOURCE      Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 662)
AUTHORS   Bonaldi,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL    Genome Res. 6 (9), 791-806 (1996)
MEDLINE    97044477
COMMENT    Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a homofide poly A
            tail. cDNA library Preparation: M.B. Soares lab Clone distribution:
            clones will be available through Research genetics (www.resgen.com)
            Seq primer: M13 forward
            POLY-A-yes.

FEATURES
Source      Location/Qualifiers
            1..662
            /organism="Rattus norvegicus"
            /strain="Sprague-Dawley"
            /db_xref="taxon:10116"
            /clone="UI-R-CNI-cj1-c-13-0-UI"
            /dev_stage="adult"
            /lab_host="DH10B (Life Technologies)"
            /note="Vector: pUT7SD-Pac (Pharmacia) with a modified
            polylinker. Site_1: Not I; Site_2: Eco RI; The UI-R-CNI
            library is a subtracted library derived from the following
            pool of seven normalized rat libraries: normalized rat
            seminal vesicles, normalized rat penis, normalized rat
            bladder, normalized rat cervix, normalized rat brown
            adipose, normalized rat fundus, and normalized rat
            salivary gland. It was constructed according to the
            procedure described by Bonaldi, Lennon & Soares (Genome
            Research Genome 6: 791-806, 1996). For construction of
            the CNI library, plasmid DNA from the pool of normalized
            libraries was electroporated into competent bacteria for
            the production of single-stranded circular DNA. This was
            then used as a tracer in a subtractive hybridization with
            a driver (PCR amplified inserts from a plasmid DNA template

```

preparation) comprising: a) a pool of about 34,000 clones from the Rat Unigene Set corresponding to plates R-5-AA-NM excluding plates R-5-MM and MN. This pool represented 40% of the final driver population. b) a pool of about 29,000 clones from subtracted libraries CA0 and CA1 corresponding to plates R-CA0-AMV through R-CA0-AXS, R-CA0-AZX through R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BJS, R-CA0-BKE, R-CA0-BKH, R-CA0-BKJ-K, R-CA0-BKP through R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA, R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through R-CA0-BOJ, R-CA0-BPA through R-CA0-BPG, R-CA1-BBA through R-CA1-BDA, R-CA1-BH2 through R-CA1-BJP, R-CA1-BJR, R-CA1-BJT through R-CA1-BKB, R-CA1-BKD, R-CA1-BKF, R-CA1-BKI, R-CA1-BKT, R-CA1-BLF, R-CA1-BLH through R-CA1-BLN, R-CA1-BLS, R-CA1-BLV, R-CA1-BNR, and R-CA1-BLE. The resulting pool represented 20% of the final driver population. c) a pool of about 15,000 clones from non-normalized libraries CS0s, CT0s, CU0s, CW0s, CX0s and normalized libraries CS0, CT0, CU0, CW0, and CX0 corresponding to plates R-CS0s-CBD through R-CS0s-CBO, R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through R-CU0s-CCA, R-CW0s-CCB through R-CW0s-CCM, R-CX0s-CCN through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through R-CS0-BTV, R-CS0-BVM, R-CT0-BTM through R-CT0-BUP, R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CW0-BVY through R-CW0-BWP, R-CW0-BXN through R-CW0-BXO, R-CX0-BWQ through R-CX0-BXM. The resulting pool represented 5% of the final driver population. d) a pool of about 5,000 clones (1,000 from non-normalized eye library CV0 and 4,000 from normalized eye library CV1) corresponding to plates R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC, R-CV1-BSE through R-CV1-BTC, and R-CV1-BVO through R-CV1-BVU. This pool represented about 5% of the final driver population. e) A pool of about 10,000 clones from subtracted library BS2, BV0 and BV0p (7-9.5 kb cDNA library fraction from rat whole embryo), and BX0 (0.5-7kb cDNA library fraction from rat whole embryo) corresponding to plates R-BS2-BDB through R-BS2-BFB, R-BV0-ANK through R-BV0-ANR, R-BV0p-AOI through R-BV0p-AOX, and R-BX0-AOY through R-BX0-ASH. The resulting pool represented 5% of the final driver population. f) a pool of about 7,000 clones from the seven non-normalized libraries that make up the tracer including CV0, CZ0, DAO, DBO, DCO, DDO, and DEO corresponding to plates R-CV0-BXP through R-CV0-BXZ, R-CZ0-BYA through R-CZ0-BYL, R-CZ0-BZB-C, R-DAO-BYJ through R-DAO-BYP, R-DAO-BZD through R-DAO-BZH, R-DBO-BYQ through R-DBO-BZA, R-DC0-BZT through R-DC0-BZQ, R-DC0-CAY through R-DC0-CBA, R-DD0-BZR through R-DD0-CAJ, R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAI. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CNO, that makes up the tracer. The corresponding plates are R-CNO-BKW through R-CNO-BLD, R-CNO-BLG, R-CNO-BLP through R-CNO-BLR, R-CNO-BLT, R-CNO-BLM-X, R-CNO-BMB, and R-CNO-BMF through R-CNO-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CNO pool corresponding to the following addresses: bkx-a-09-0-UI, bkx-b-09-0-UI, bkx-b-11-0-UI, bkx-b-10-0-UI, bkx-d-01-0-UI, bkx-d-06-0-UI, bkx-g-08-0-UI, bkx-h-12-0-UI, bkx-a-05-0-UI, bkx-a-06-0-UI, bkx-a-11-0-UI, bkx-c-09-0-UI, bkx-d-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, bla-b-12-0-UI, bla-f-02-0-UI, bla-a-11-0-UI, bla-e-95-0-UI, bla-f-08-0-UI, bla-f-02-0-UI, bla-h-04-0-UI, bla-a-05-0-UI, bla-f-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CNO clone (corresponding to the address bkx-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG_SEQ=None found"



GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:52:54 ; Search time 90 Seconds
(without alignments)
879.134 Million cell updates/sec

Title: US-09-937-060A-5
Perfect score: 2016
Sequence: 1 MDPAGGPRGVLPKPCRVLL.....CPEPPSWKPOQMPPEEPL 384

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mhc:.*
8: sp_organelle:.*
9: sp_phage:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*
15: sp_virus:.*
16: sp_bacteria:.*
17: sp_archaeap:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2009	99.7	384	4	Q96HV8
2	1999	99.2	398	4	Q96GK1
3	1644	81.5	329	6	Q9N0A5
4	1630.5	80.9	381	11	Q912N3
5	1627.5	80.7	388	11	Q912N3
6	1627.5	80.7	388	11	Q912N3
7	1599.5	79.3	383	11	Q912N3
8	1508	74.8	290	4	Q9B7G7
9	886.5	44.0	634	4	Q9H0Q2
10	886.5	44.0	634	4	Q9H0Q2
11	608	30.2	384	11	Q9D4D1
12	562	27.9	641	5	Q9VYH8
13	521.5	25.9	907	5	Q9VZM0
14	521	25.8	661	5	Q8T0C1
15	451	22.4	1240	10	Q65419
16	411	20.4	624	5	Q8T269

17	374.5	18.6	473	5	Q18425	Q18425 caenorhabdi
18	363.5	18.0	354	4	Q9NMU7	Q9NMU7 homo sapien
19	356.5	17.7	687	3	Q06147	Q06147 saccharomyc
20	355	17.6	458	3	Q14159	Q14159 schizosacch
21	345	17.1	624	3	Q12246	Q12246 saccharomyc
22	272	13.5	763	10	Q9LRB0	Q9LRB0 arabidopsis
23	265	13.1	757	10	Q94HY9	Q94HY9 oryza sativ
24	264.5	13.1	481	4	Q9BYH3	Q9BYH3 homo sapien
25	264.5	13.1	537	4	Q8TCM0	Q8TCM0 homo sapien
26	250.5	12.4	549	5	Q9TZ11	Q9TZ11 caenorhabdi
27	250	12.4	732	10	Q9FHL3	Q9FHL3 arabidopsis
28	239	11.9	487	5	Q95S15	Q95S15 drosophila
29	239	11.9	596	5	Q9VNA6	Q9VNA6 drosophila
30	227.5	11.3	326	4	Q9UGF5	Q9UGF5 homo sapien
31	213.5	10.6	422	4	Q9NPA8	Q9NPA8 homo sapien
32	203	10.1	421	11	Q9ESW4	Q9ESW4 mus musculu
33	185	9.2	533	10	Q9LU45	Q9LU45 arabidopsis
34	162.5	8.1	406	5	Q9VLJ0	Q9VLJ0 drosophila
35	155.5	7.7	306	16	Q927W6	Q927W6 listeria in
36	148	7.3	173	11	Q9D087	Q9D087 mus musculu
37	140.5	7.0	306	16	Q8Y497	Q8Y497 listeria mo
38	140.5	7.0	338	10	Q94C71	Q94C71 arabidopsis
39	136.5	6.8	309	16	Q34799	Q34799 bacillus su
40	131	6.5	586	10	Q949C3	Q949C3 oryza sativ
41	130	6.4	295	16	Q9KBM4	Q9KBM4 bacillus ha
42	129.5	6.4	303	16	Q31502	Q31502 bacillus su
43	127	6.3	342	16	Q9CTE3	Q9CTE3 lactococcus
44	126	6.2	302	16	Q9CEN8	Q9CEN8 lactococcus
45	125	6.2	340	16	Q9A0J4	Q9A0J4 streptococc

ALIGNMENTS

RESULT 1	ID	Q96HV8	PRELIMINARY:	PRT:	384 AA.
AC	Q96HV8:				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)				
DE	Unknown (protein for MGC:15041).				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=SKIN;				
RA	Strausberg R.;				
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC008040; AAH08040.1; -				
DR	InterPro: IPR001395; Aldo/kel_red.				
DR	InterPro: IPR001206; DAGC.				
DR	InterPro: IPR003622; DAG_kin_cat.				
DR	Pfam: PF00781; DAGKc. 1.				
DR	ProDom: PD005043; DAG_kin_cat. 1.				
DR	PROSITE; PS00063; ALDOKE10_REDUCTASE_3; UNKNOWN_1.				
SO	SEQUENCE 384 AA; 42474 MW; P82999F306113B0 CRC64;				
Query Match		99.7%;	Score 2009;	DB 4;	Length 384;
Best Local Similarity		99.7%;	Pred. No. 5e-168;		
Matches 383;	Conservative	0;	Mismatches 1;	Indels	0;
QY	1	MDPAGGPRGVLPKPCRVLLNPRGKGKALQLFSSHVQPLLAETISFTLMLTRRNA	60		
DB	1	MDPAGGPRGVLPKPCRVLLNPRGKGKALQLFSSHVQPLLAETISFTLMLTRRNA	60		
QY	61	RELVSEELGRMDALVNVSGGLMHEVYVNGLMERDMETAIQKPLCSIPASSGNALASL	120		
DB	61	RELVSEELGRMDALVNVSGGLMHEVYVNGLMERDMETAIQKPLCSIPASSGNALASL	120		
QY	121	NHYAGVEQVTNEDLLTNCILLICRRLLSPMNLLSHFTASGLRFLSVLSIANGFIADVDLE	180		

```

Db 121 NHYAGBOVTNEDLLTNCITLLCRRLSPNNLSLHTASGLRLEFSVLSLAWGFIADVDLE 180
QY 181 SEKYRRLGEMRFTLGTFLRLAALRTYRGRLAYLPVGRVSGSKTPASPVVVOGDPVDAHLVP 240
Db 181 SEKYRRLGEMRFTLGTFLRLAALRTYRGRLAYLPVGRVSGSKTPASPVVVOGDPVDAHLVP 240
QY 241 LEEPPVSHMTVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVWHLFFYVRAGVSRAML 300
Db 241 LEEPPVSHMTVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVWHLFFYVRAGVSRAML 300
QY 301 RLFLAMEKGRHMEYECPLVYVVAFRLEPKDKGVAVNGELMVSEAVOGVHPNYFM 360
Db 301 RLFLAMEKGRHMEYECPLVYVVAFRLEPKDKGVAVNGELMVSEAVOGVHPNYFM 360
QY 361 MYSGCYEPPPSKPKQMPPEEP 384
Db 361 MYSGCYEPPPSKPKQMPPEEP 384

```

RESULT 2

```

ID 096CK1 PRELIMINARY: PRT: 398 AA.
AC 096CK1:
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
DE Unknown (protein for MGC:15040).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RA TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC009419; AAI09419.1;
DR InterPro: IPR001395; Aldo/Ket_red.
DR InterPro: IPR001206; DAGKC.
DR InterPro: IPR003622; DAG_Kin_cat.
DR Pfam: PF00781; DAGKC; 1.
DR ProDom: PD005043; DAG_Kin_cat; 1.
DR PROSITE: PS00063; ALDOXETO_REDUCTASE; 3; UNKNOWN_1.
SO SEQUENCE 398 AA; 43944 MW; 7F2C9B26C030E560 CRC64;

```

Query Match 99.2%; Score 1999; DB 4; Length 398;

Best Local Similarity 96.5%; Pired. No. 3, 9e-107;

Matches 384; Conservative 0; Mismatches 0; Indels 14; Gaps 1;

```

QY 1 MDP-----ACGRCVLPKPCRVLYLNPGRGKALQLFPSHYOPLLAEE 46
Db 1 MDPVVGCGRGLEGFVFSAGRGVLPKPCRVLYLNPGRGKALQLFPSHYOPLLAEE 60
QY 47 ISFTLMLTERRNHARELVRSEBELGRMDALVYVSGDGLMHEVYVNGLMERPMEALOKPCL 106
Db 61 ISFTLMLTERRNHARELVRSEBELGRMDALVYVSGDGLMHEVYVNGLMERPMEALOKPCL 120
QY 107 SLPAGSGNALAASLNHYAGYEOVTNEDLLTNCITLLCRRLSPNNLSLHTASGLRLEFSV 166
Db 121 SLPAGSGNALAASLNHYAGYEOVTNEDLLTNCITLLCRRLSPNNLSLHTASGLRLEFSV 180
QY 167 LSLANGFIADVDLESEKRRLGEMRFTLGTFLRLAALRTYRGRLAYLPVGRVSGSKTPASP 226
Db 181 LSLANGFIADVDLESEKRRLGEMRFTLGTFLRLAALRTYRGRLAYLPVGRVSGSKTPASP 240
QY 227 VYVVOGDPVDAHLVPLEPPVSHMTVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVWHL 286
Db 241 VYVVOGDPVDAHLVPLEPPVSHMTVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVWHL 300
QY 287 LTYVRAGVSRAMLRLFLAMEKGRHMEYECPLVYVVAFRLEPKDKGVAVNGELMVSEAV 346
Db 301 LTYVRAGVSRAMLRLFLAMEKGRHMEYECPLVYVVAFRLEPKDKGVAVNGELMVSEAV 360

```

```

QY 347 SEAVOGVHPNYFMVWVSGCYEPPPSKPKQMPPEEP 384
Db 361 SEAVOGVHPNYFMVWVSGCYEPPPSKPKQMPPEEP 398

```

RESULT 3

```

ID 09N0A5 PRELIMINARY: PRT: 329 AA.
AC 09N0A5:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
DE Unnamed protein product.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN 11
RP SEQUENCE FROM N.A.
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB046025; BAB01607.1;
DR InterPro: IPR001395; Aldo/Ket_red.
DR InterPro: IPR001206; DAGKC.
DR InterPro: IPR003622; DAG_Kin_cat.
DR Pfam: PF00781; DAGKC; 1.
DR ProDom: PD005043; DAG_Kin_cat; 1.
DR SMART: SM00046; DAGKC; 1.
DR PROSITE: PS00063; ALDOXETO_REDUCTASE; 3; UNKNOWN_1.
SO SEQUENCE 329 AA; 36393 MW; 7F430AB0CA5FDC7 CRC64;

```

Query Match 81.5%; Score 1644; DB 6; Length 329;

Best Local Similarity 98.1%; Pired. No. 4, 4e-136;

Matches 316; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MDPAGGPGVLPKPCRVLYLNPGRGKALQLFPSHYOPLLAESTLWITERRNHA 60
Db 1 MDPAGGPGVLPKPCRVLYLNPGRGKALQLFPSHYOPLLAESTLWITERRNHA 60
QY 61 RELVSEBELGRMDALVYVSGDGLMHEVYVNGLMERPMEALOKPCLSPAGSGNALASL 120
Db 61 RELVSEBELGRMDALVYVSGDGLMHEVYVNGLMERPMEALOKPCLSPAGSGNALASL 120
QY 121 NHYAGYEOVTNEDLLTNCITLLCRRLSPNNLSLHTASGLRLEFSVLSLAWGFIADVDLE 180
Db 121 NHYAGYEOVTNEDLLTNCITLLCRRLSPNNLSLHTASGLRLEFSVLSLAWGFIADVDLE 180
QY 181 SEKYRRLGEMRFTLGTFLRLAALRTYRGRLAYLPVGRVSGSKTPASPVVVOGDPVDAHLVP 240
Db 181 SEKYRRLGEMRFTLGTFLRLAALRTYRGRLAYLPVGRVSGSKTPASPVVVOGDPVDAHLVP 240
QY 241 LEEPPVSHMTVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVWHLFFYVRAGVSRAML 300
Db 241 LEEPPVSHMTVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVWHLFFYVRAGVSRAML 300
QY 301 RLFLAMEKGRHMEYECPLVYVVAFRLEPKDKGVAVNGELMVSEAVOGVHPNYFM 360
Db 301 RLFLAMEKGRHMEYECPLVYVVAFRLEPKDKGVAVNGELMVSEAVOGVHPNYFM 360

```

RESULT 4

```

ID 0912N3 PRELIMINARY: PRT: 381 AA.
AC 0912N3:
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Sphingosine kinase 1a.

```


Best Local Similarity 81.4%; Pred. No. 2,1e-134;
Matches 307; Conservative 33; Mismatches 36; Indels 1; Gaps 1;

```
QY 7 PRGVLPKRVYLVLLNPGCGKALQIFRSHVOPPLAEAEISFTLTERNRHARELVRS 66
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 129 PRGLPRPCRVLYLLNPGCGKALQIFRSVOPPLAEAEITFTLTERKNHARELVCA 188
QY 67 EELGRMDALVMSGDGLMHEVNGLMERPDMEATOKPLCSLPGSGNALAASLVNHAYG 126
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 189 EELGRMDALVMSGDGLMHEVNGLMERPDMEATOKPLCSLPGSGNALAASLVNHAYG 248
QY 127 EGYTNEEDLLNCTLLCRRLSPMNLSTHTASGLRFSYSLMAGFLAVDSEKRYR 186
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 249 EGYTNEEDLLNCTLLCRRLSPMNLSTHTASGLRFSYSLMAGFLAVDSEKRYR 308
QY 187 LCGEMFTLGTFLRLAALRTYRGLATLPVGRVSGKTPASPVYVQGGPYDAHLVPLEPVP 246
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 309 LGEIRFTVGTFFRLASLRTYOGQLATLPVGTASKRPAS-FLVQKGPVDTLHVLPLEPVP 367
QY 247 SHMTVVPDEFLVYLVALLSHSGSEMFAPMGRCAGVYMHLEFYRAGVSRAMLRLFLAM 306
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 368 SHMTVVPDEFLVYLVALLSHSGSEMFAPMGRCAGVYMHLEFYRAGVSRAMLRLFLAM 427
QY 307 EKGRIHMEYECPLYVYVVAFRLEPKDKGVFAVDGLMVSBAVQGVHNPYFMVSGCV 366
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 428 OKGKHHELDCLPLVYVVAFRLEPKDKGVFAVDGLMVSBAVQGVHNPYFMVSGCV 487
QY 367 EPPPSWKPOQMPPEEP 383
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 488 DAPSGRSDRSGPPEEP 504
```

RESULT 7

Q91V26 PRELIMINARY; PRT: 383 AA.
AC 091V26:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Sphingosine kinase 1f (Sphingosine kinase 1a) (Sphingosine kinase 1c)
GN SPHK1F OR SPHK1C OR SPHK1D OR SPHK1E.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Imamura T., Ohgane J., Ito S., Ogawa T., Hattori N., Tanaka S.,
RA Shiohara K.;
RT "Cpg island of rat sphingosine kinase-1 gene: tissue-dependent DNA
RT methylation status and multiple alternative first exons.";
RL Genomics 78:117-125(2001).
DR EMBL: AB049575; BAB62324.1; -
DR EMBL: AB049575; BAB62320.1; -
DR EMBL: AB049575; BAB62321.1; -
DR EMBL: AB049575; BAB62322.1; -
DR EMBL: AB049575; BAB62323.1; -
DR EMBL: AB049575; BAB62323.1; -
DR InterPro: IPR001206; DAGC.
DR InterPro: IPR003622; DAGC.
DR Pfam: PF00781; DAGC.
DR ProDom: PD005043; DAGC.
KM kinase.
SQ SEQUENCE 383 AA; 42418 MW; CAAD5817BF4B9507 CRC64;

Query Match 79.3%; Score 1599.5; DB 11; Length 383;
Best Local Similarity 78.1%; Pred. No. 4.3e-132;
Matches 300; Conservative 42; Mismatches 41; Indels 1; Gaps 1;

```
QY 1 MDPAGCPRCVLPKRVYLVLLNPGCGKALQIFRSHVOPPLAEAEISFTLTERNRHA 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MDPADCPKRLPLPRCVLYLVLLNPGCGKALQIFRSVOPPLAEAEISFTLTERNRHA 60
```

QY 61 RELVASELGRMDALVMSGDGLMHEVNGLMERPDMEATOKPLCSLPGSGNALAASL 120
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 RELVASELGRMDALVMSGDGLMHEVNGLMERPDMEATOKPLCSLPGSGNALAASL 120

```
QY 121 NHVAGYDGVNEDLLNCTLLCRRLSPMNLSTHTASGLRFSYSLMAGFLAVDSE 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 NHVAGYDGVNEDLLNCTLLCRRLSPMNLSTHTASGLRFSYSLMAGFLAVDSE 180
QY 181 SEXYRRLGEMFTLGTFLRLAALRTYRGLATLPVGRVSGKTPASPVYVQGGPYDA 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 SEXYRRLGEMFTLGTFLRLAALRTYRGLATLPVGRVSGKTPASPVYVQGGPYDA 240
QY 241 LEEYPSLWYVYVDEDFVYLVALLSHSGSEMFAPMGRCAGVYMHLEFYRAGVSR 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 LEEYPSLWYVYVDEDFVYLVALLSHSGSEMFAPMGRCAGVYMHLEFYRAGVSR 300
QY 301 RLFLAMEKGRHMEYECPLYVYVVAFRLEPKDKGVFAVDGLMVSBAVQGVHNPYFM 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 RLFLAMEKGRHMEYECPLYVYVVAFRLEPKDKGVFAVDGLMVSBAVQGVHNPYFM 360
QY 361 MVSGCVPPEPPSWKPOQMPPEEP 384
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 MVSGCSDPSGCRDSDORPPEEP 383
```

RESULT 8

Q9BTG7 PRELIMINARY; PRT: 290 AA.
AC 09BTG7:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Similar to sphingosine kinase 1 (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue-Skin;
RC Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC004112; AA04112.1; -
KW Kinase.
FT NON_TER
SQ SEQUENCE 290 AA; 31998 MW; A3B7A219DB52C520 CRC64;

Query Match 74.8%; Score 1508; DB 4; Length 290;
Best Local Similarity 100.0%; Pred. No. 3.2e-124;
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 98 EFAIOLKPLCSLPGSGNALAASLVNHAYGBOVTNEDLLTNCCTLLCRRLSPMNLSTHT 157
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 EFAIOLKPLCSLPGSGNALAASLVNHAYGBOVTNEDLLTNCCTLLCRRLSPMNLSTHT 63
QY 158 ASGLRLFSYSLMAGFLAVDSEKRYRLEGEMFTLGTFLRLAALRTYRGLATLPVGR 217
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 ASGLRLFSYSLMAGFLAVDSEKRYRLEGEMFTLGTFLRLAALRTYRGLATLPVGR 123
QY 218 VGSKTPASPVYVQGGPYDAHLVPLEPVSMTVYVDEDFVYLVALLSHSGSEMFAPM 277
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 124 VGSKTPASPVYVQGGPYDAHLVPLEPVSMTVYVDEDFVYLVALLSHSGSEMFAPM 183
QY 278 GRCAAGVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 337
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 GRCAAGVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 243
QY 338 FAVDGLMVSBAVQGVHNPYFMVSGCVPEPPSWKPOQMPPEEP 384
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 244 FAVDGLMVSBAVQGVHNPYFMVSGCVPEPPSWKPOQMPPEEP 290
```

RESULT 9

ID	Query Match	Best Local Similarity	Matches 195; Conservative	Score 44.0%; 39.8%;	DB 4; Pred. No. 3.1e-69;	Length 654;	Mismatches 107;	Indels 127;	Gaps
09H002	PRELIMINARY;								
09H002	PR1: 654 AA.								
AC	09H002.								
DT	01-MAR-2001 (TREMBLrel. 16, Created)								
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)								
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)								
DE	Hypothetical 69.2 kDa protein.								
CN	DKF2564M0682.								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.								
OX	NCBI_taxid:9606;								
RN	11								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=BRIN.								
FX	MEDLINE-21154917; PubMed-11230166;								
RA	Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,								
RA	Ansober W., Boecker M., Bloeker H., Bauersachs S., Blum H.,								
RA	Lauber J., Duesterhoef A., Beyer A., Koehrer K., Strack N.,								
RA	Mewes H.W., Oltmawelder B., Obermayer B., Tampe J., Heubner D.,								
RA	Manubut R., Korn B., Klein M., Poustka A.;								
RT	"Towards a Catalog of Human Genes and Proteins: Sequencing and								
RT	Analysis of 500 Novel Complete Protein Coding Human CDNs."								
RL	Genome Res. 11:422-435(2001).								
DR	EMBL; AL136701; CAB66636.1; -								
DR	InterPro: IPR001206; DAGC.								
DR	InterPro: IPR003622; DAG_kin_cat.								
DR	Pfam; PF00781; DAGC.								
DR	ProDom; PD005043; DAG_kin_cat; 1.								
DR	SMART; SM00046; DAGC; 1.								
KW	Hypothetical protein.								
SO	SEQUENCE 654 AA; 69207 MW; E9DF12E2AC0CF0FB CRC64;								
0Y	Query Match	44.0%; Score 886.5; DB 4; Length 654;							
0Y	Best Local Similarity	39.8%; Pred. No. 3.1e-69;							
0Y	Matches 195; Conservative	61; Mismatches 107; Indels 127; Gaps							
0Y	10 VPRRCRVYLINPFGGKGKALQFRSHVQPLAEFISFTMLERRNHARELRESEL 69								
DB	176 ILPRPRRLILVNPFGGKGLAMQCKKNHVLPISEGLSFLNLTQRQNHARELVQGLSL 235								
0Y	70 GMDALVYMSGDGLMHEVYNGLMERPMDETAIOKPLCSLPAGSGNLAASLHHVAGYQV 129								
DB	236 SEMDGIIVVSSDGLLHEVYNGLLDRPDMEAVKMPVGIPLCGSGNALAGAVYQHGFEPA 295								
0Y	130 TNEEDLTNCTLLCRLLSPANLSLHTASGRLTSVLSLANGFTADVDSSEKRRRGE 189								
DB	296 LGLDILNCSLILKGGGGLPDLISVTLASGRCSFSLVAMGFVSDVDIOSERRFALGS 355								
0Y	190 MRFYTGFLRLAALRTYRGRLAYLPVG-RVGSKTPA----- 224								
DB	356 ARFTLGTVGLATLTHTYRGRLSYLTAVERPASPTFAHSLPRAKSELTLTPDPAPPAHASP 415								
0Y	225 ----- 224								
DB	416 LHRVSVDLPPLPQPALASPGSPPEPLILSLNGGSGPELAGDWGAGADAPLSPDLLSSP 475								
0Y	225 -----SPVYVQO-----CPVNAHLVPLEEPPSPSHWTVP 253								
DB	476 GSPKALHSPVSEGAADVIPPSSGLPLPPDPARVAGASTCGPPHLLPPLTPPDM-VTL 534								
0Y	254 DDDPFLVTLALHSHSGSEMFAPMGRCAQVNAHLYFVRAVSRAMLLRLFLAMEGRHME 313								
DB	535 EDDFVYLMALISPSHIGADLVAPARFDDGLVHLCLVRSGLISRALTLRLFLAMERSHFS 594								
0Y	314 YCCPVLVYVPAVRPKDGKVAVDGELNAVSEAVQGVHPNFMVSGCVERPPSK 373								
DB	595 LCCPOLGVAAARAFRIEPLTPRGVILTVDEQVEYGLAQMHQHPGIGTLTLTG----PPGC- 649								
0Y	374 PQQMPPEEP 383								
DB	650 -----PGRPEP 654								

RESULT	10		
Q9BRN1			
ID	Q9BRN1	PRELIMINARY:	PRT: 654 AA.
AC	Q9BRN1:		
DT	01-JUN-2001	(TREMBLrel. 17, Created)	
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)	
DE	Hypothetical 69.2 kda protein.		
OS	Homo sapiens (human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Lymph;		
RL	Strausberg R.;		
RL	Submitted (APR-2001) to the EMBL/Genbank/DBDJ databases.		
DR	EMBL: BC006161; AAH06161.1; -		
DR	InterPro: IPR001206; DAgKC:		
DR	InterPro: IPR003622; DAg_Kin_cat.		
DR	Pfam: PF00781; DAgKC: 1.		
DR	ProDom: PD05043; DAg_Kin_cat; 1.		
DR	SMART: SM00046; DAgKC: 1.		
KW	Hypothetical protein.		
SO	SEQUENCE 654 AA; 69217 MW; F73FFCEC930DA50F CRC64;		

Query Match Similarity	44.0%	Score 886.5	DB 4:	Length 654
Best Local Similarity	39.8%	Pred. No. 3.1e-69		
Matches 195:	Conservative	61:	Mismatches 107:	Indels 127:
Gaps				
6				

```

DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Adult male testis cDNA, RIKEN full-length enriched library.
DE Clone:4933402E20, full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shigaawa A., Shibata K., Yoshino M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glass C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiri L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bolunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Segal T., Shibata T., Storch K.-F.,
RA Suzuki H., Togo O.K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Functional annotation of a full-length mouse cDNA collection."
RL NCBI:403:685-690(2001).
DR EMBL: AK016616; BAB0339.1;
SQ SEQUENCE 384 AA: 39865 MW: 6C94D1D3806092A3 CRC64:

Query Match 30.2%; Score 608; DB 11; Length 384;
Best Local Similarity 36.4%; Pred. No. 4, 3e-45;
Matches 141; Conservative 38; Mismatches 88; Indels 120; Gaps 5;

OY 104 PLCSLPGAGSNAALASLNHYAGYEQVNTEDLTNCTLLCRRLSPNNLSLHTASGLRL 163
DB 2 PLGVLPGSSGALALGAVNHNGFEQVGVLDLNCSLICSGSPRLDLISVTLASGRC 61
OY 164 FSVLSLAWGFIDVDVLESEKRYRSGEMKFTLGPRLALRTYRGRLATLVY----- 215
DB 62 FSVLSLAWGFIDVDVLESEKRYRSGEMKFTLGPRLALRTYRGRLATLVY----- 215
OY 216 ----- 215
DB 122 PGHSPLRAKSELVLAAPAPAAATHTSPHSVSDLPRLPQALVSPGSPRLPDLISLNG 181
OY 216 -----GRGS-----KTFAS-----PVY 228
DB 182 GRELTGDMGAGDAPLSPDLSPSSPNALKTAQSLFIAEGPREMPASGSLPRTHSAPR 241
OY 229 VQGGQVDAHLVPLEBVPVPSHTVTVDDPEVYVLAHLSHLSGEMFADMGSCAAGVNL 288
DB 242 STWGVNDLRLPLSPRLPQDMVTIEE-FVLMIGLIPSHCADLMAAFAHAFDGVVNL 300
OY 289 YVAGVSRAMLRLFLMEKGRHMEYECPLVYVYVAFRLPKDKGVAFVDELVSE 348
DB 301 VWRSGSRALLRLTLMEHNGHNSLSCPHLGAARAFRLERPLRGLLTVDELVEVS 360
OY 349 AVQGVNPRFMVNSGCVPEPPSKPQ 375
DB 361 PLOAVNHGLATLTLTG---PAGCKPQ 383

```

```

DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DE 01-OCT-2001 (TREMBLrel. 18, last annotation update)
DE CG1747 protein.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amandus P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Bakendole J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Brinkman P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA de Fabros B., Delcher A., Deng Z., Davis A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,
RA Foster A., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harits N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mount S.M., Moy M., Murphy B., Murphy C., Morris J., Mostrefi A.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M.,
RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relier K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spiet E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Switskas R., Tector C., Turner R., Ventler E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Glods R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster."
SCIENCE 287:2185-2193(2000).
DR EMBL: AE003486; AF48045.1;
DR FLYBASE: FBgn0030300; Drosophila CG1747.
DR InterPro: IPR001206; DAGC.
DR Pfam: PF00781; DAGC_kin_cat.
DR ProDom: PD005043; DAGC_kin_cat; 1.
DR SMART: SM00046; DAGC_kin_cat; 1.
SQ SEQUENCE 641 AA: 71421 MW: 1CE24E480F14758D CRC64:

Query Match 27.9%; Score 562; DB 5; Length 641;
Best Local Similarity 32.5%; Pred. No. 9, 1e-41;
Matches 148; Conservative 67; Mismatches 121; Indels 120; Gaps 16;

OY 12 PRPC--RVVLVLPNPGKGAALQFPRSHVQPLLAFAISFTMLTERNRNARLVSEF 68
DB 184 PADCGKQLILLNPKSSGKGRLEFQVAPRLTAAYOVYDOLITTPVAKFEVTRRD 243
OY 69 -LGRNDALVNSGGLMHEVYVNCILMERPRMFAIOK-PGLSLPGAGSNAALASLNHYAG- 125
DB 244 LTRYSGLTVASGGLPEYVNLGLMERMDWRACRELPLGLITPGSSNGSLAKSVANHNE 303
OY 126 -YEQVNTEDLTNCTLLCRRLSPNNLSLHTASGLR--LFSVLSLAWGFIADVLES 181

```

```

Db 304 PYERKP-----LHATLTCMAKCKSTPMDDVVELATROKHFWYSLVSGWGLADIIDIES 359
QY 182 EKYRILGEMRFTLTFTLALRTYRGLAVLPVGRVSKTPASPV----- 227
Db 360 ERLSIGMORTLWAKRLIGLRKGVSTL-LGK-GKKEP--PVEARELPASTAAG 415
QY 228 -----VVVO----- 231
Db 416 IRSSLPLNAGEFHDLPEEEGEAVLDGEQADATSLDRSVYRQHADSMSMRTAYVS 475
QY 232 --GP-----VDA-----HLVPLEEVPSSH--WTVPVDEDFVLYL 261
Db 476 LGGPSMRSNRSMSISORIEANAEFAERVPPTGTPLOMPDLSSDGM-ICEEDGDFVAVH 534
QY 262 ALLSHSGEMFAAPMGRCAGVAMHLPVVRAGVRAMLRLFLAMEKGRHMEY-ECPTLYV 320
Db 535 AAYTHSSSDVFAPESLDLGLYLTIRKGVSRHOLLNMLNMACTHLP.IGEDPTIK 594
QY 321 YVPVAFRLERPKDGKGVFAVDGELVSEAVOGVHP 356
Db 595 VVPCAFRIEPPSSDGLVYDGEREYGPICAFVHP 630

RESULT 13
QYVZWO PRELIMINARY: PRT: 907 AA.
AC QYVZWO:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE CG2159 protein.
GN CG2159.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Aamatiides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George G.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris W.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Honck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jajali M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclele J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Schelder F., Shen H.,
RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Ventler E., Wang A.H., Wang X.,

```

```

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003476; AAF47706.1; -.
DR Flybase; FBgn0035391; CG2159.
DR InterPro; IPR001251; CRAL_TRIO.
DR InterPro; IPR001206; DAGKC.
DR InterPro; IPR003622; DAG_Kin_cat.
DR Pfam; PF00650; CRAL_TRIO; 1.
DR Pfam; PF00781; DAGKC; 1.
DR ProDom; PD005043; DAG_Kin_cat; 1.
DR SMART; SM00046; DAGKC; 1.
DR SMART; SM00516; SEC14; 1.
SO SEQUENCE 907 AA; 102561 MW; 58206C8E12A3F195 CRC64;

Query Match 25.9%; Score 521.5; DB 5; Length 907;
Best Local Similarity 30.5%; Pred. No. 5,1e-37;
Matches 134; Conservative 61; Mismatches 138; Indels 107; Gaps 8;

QY 16 RVLVLNPRGKGKALQLFRSHVQPLAEAEISFTLMLTERNNHARELYRSEELGRDAL 75
Db 216 RVLVLNPKSGSGDAREVFNHVTPLNEAEVYDLYTKHSNFAIEFLSTRCLDAMCCV 275
QY 76 VYMSDGLMHEVNVGLMRRPMEYAI-QKPLSLPAGSGNLAALNINAYGEQYTNEDL 134
Db 276 VAVGDDGFEHEIVNGLLORODAHVLPILALGIIIPCGSGNGIARSIAH-----CYNRPV 329
QY 135 LTNCITLLCRLLSPMNLISLHTASGLRLEFSVLSAMGFADVDESEKRYRLGEMRFTL 194
Db 330 LGALTYSIGR-SSPMYVRYVQDS-RSLYSFLSLIGWGLISVODLESIRIKLGIQRTFV 387
QY 195 GTFLALALRTYRGLAVLPVGRVSKTPASPVVVOGCV-----DAHLVLEEPVSH 248
Db 388 WLYRLVNLRTYNGRISYLLFTDHEVSSYTHSAGVAAORMOSSRSCNTHIDMLNGPAPLY 447
QY 249 -----WT----- 250
Db 448 HSSAEYLPQEFADVISLETSLNQSFRSRSQSVLSGSGSRSPYSISSELYSLADESEFA 507
QY 251 -----VYDEDFVVLALLSHSGEMFAAPM 277
Db 508 GLAASLNRQONTGPASRLPPLNPLESEDQWLYEKEGFVMAHAYQTIGIDCHFPAPK 567
QY 278 GRCAGVAMHLPVVRAGVRAMLRLFLAMEKGRHMEY-ECPTLYVVPVAFRLERPKDGK 336
Db 568 AOLNCGTYLILIRAGISPHLLSFLYNNMSGTHLPESHDDHVKLVPVAFRLERPDYDHG 627
QY 337 VFAVDGELVSEAVOGVHP 356
Db 628 IITVDGERVEFPGLOAEVLP 647

RESULT 14
QYVZWO PRELIMINARY: PRT: 661 AA.
AC QYVZWO:
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE LD11247P.
GN CG2159.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;

```

RA Stapleton M., Brockstein P., Hong L., Agbayani A., Carlson J.,
RA Change M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacheco J., Paragans V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celisner S.,
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY069417; AAL39562.1;
SQ SEQUENCE 661 AA: 73870 MW: 4266E191EBF82D08 CRC64:

Query Match 25.88: Score 521; DB 5; Length 661;
Best Local Similarity 30.28; Pred. No. 3.8e-37;
Matches 133; Conservative 60; Mismatches 145; Indels 102; Gaps 7;

OY 16 RYLYLNPBGCKALQIFRSHVOPPLAEAEISFTLMTERRNARELVSEELGRMDL 75
DB 216 RYLYLNPBGCKALQIFRSHVOPPLAEAEISFTLMTERRNARELVSEELGRMDL 75
OY 76 VYVSGDGLMEYVNGIMERPDMETAI-OKPLCSIPAGSGNALASLNHYAGYEQVYNEDL 134
DB 276 VYVSGDGLMEYVNGIMERPDMETAI-OKPLCSIPAGSGNALASLNHYAGYEQVYNEDL 134
OY 135 LTRCTLLCRRLSPMNLISHTASGLRFSVLSAMGFADVDLESEKTRRIGEMFTL 194
DB 334 VLGALITYSGSSPMADYKQLOD-RSLTSLTSGMGLSDYDIESRITMLDQRETY 392
OY 195 GTFELALFTYRGRLATPVGRVSGTASPVYVQGGV-----DAHLVPLEEPVPSH 248
DB 393 WLYLYRVLNFTYNGRISLTLTDHEVSTSHATGYAARMOSSRSCNHIIMLNGPAPLY 452
OY 249 -----WT----- 250
DB 453 HSSAEVLPQEFADVTSLETSTINOSFRSCDSMLSGSRSTFYSTISSTYSLADESEFA 512
OY 251 -----VYVDEPFVYVYALHSHSGSMFAPM 277
DB 513 GLAASLENRQONTYGPASLEPDNEPSEDOGWLYDEGEFVMAHAYOTHTGIDCHAPK 572
OY 278 GRAGVNHILFVYVAGSRAMLRFLAMEKGRHM-ETVCPYLYVYVPAARLPKPKGK 336
DB 573 AQINDGTITLIRAGISRPILHSFLYNNSSGTHLPESHDDHVKVLPVAPARLPKPKGK 632
OY 337 VFAVDELMVSPAVOGGYVP 356
DB 633 IIVDGERVEFGPIQAEVLP 652

RESULT 15
ID 065419 PRELIMINARY: PRT: 1240 AA.
AC 065419:
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Hypothetical 138.7 kDa protein.
GN F18E5.160 OR AT4G21540.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidops.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;
RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL022603; CAI18718.1;
DR EMBL: AL161555; CAB81261.1;
DR InterPro: IPR001206; DAKG.
DR InterPro: IPR003622; DAG_Kin_cat.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00781; DAKG; 2.
DR ProDom: P0005043; DAG_Kin_cat; 2.
DR SMART: SM00046; DAKG; 2.
DR PROSITE: PS00290; Ig_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 1240 AA: 138683 MW: 3AB1FED5CE099E CRC64:

Query Match 22.48: Score 451; DB 10; Length 1240;
Best Local Similarity 31.98; Pred. No. 1.2e-30;
Matches 122; Conservative 63; Mismatches 127; Indels 70; Gaps 13;

OY 11 LPRCPVLYLNPBGCKALQIFRSHVOPPLAEAEISFTLMTERRNARELVSEELG 70
DB 374 LPRCPVLYLNPBGCKALQIFRSHVOPPLAEAEISFTLMTERRNARELVSEELG 70
OY 71 RMDALVYVSGDGLMEYVNGIMERPDMETAIOKPLCSIPAGSGNALASLNHYAGYEQV 130
DB 434 KIDGIVCYSGDGLMEYVNGIMERPDMETAIOKPLCSIPAGSGNALASLNHYAGYEQV 130
OY 131 NEDTLTCTLLCRRLSPMNLISHTASGLRFSVLSAMGFADVDLESEKTRRIGEM 190
DB 494 N-----SATSIIFGHKRSYDVATI-AQCNKTFYSVLLAMGLADIDISEKFRMGSA 547
OY 191 RETLGTGL-----RLAALFTYRGRLATPV-GRVSGTASPVYVQ- 230
DB 548 RIDFVYVCLYVDFDYWCYLAIVKLLALORITCLRYNGRLIFLPAGFGYCGOPASCLYGE 607
OY 231 -----OGP-----VDALVPLEEPVSHVYVPEDEPVYVYALHSHSGSMFAPM 274
DB 608 PHSVDEKGYCGPETFREDELEMRMKGPFTITM-----LHVNPAGSNTLT 653
OY 275 APMGCAAGVNHILFVYVAGSRAMLRFLAMEKGRHM-ETVCPYLYVYVPAARLPKPKG 329
DB 654 APAAFSGYGLDLVYLR-NCPRVYVLSLMROTSSGTHV--ESPFIYIKITVEKVAFL 710
OY 330 EP-----KDGKGFVAVDGLM 345
DB 711 EPGALVDEPDKEGILDSGEVL 732

Search completed: May 9, 2003, 16:56:56
Job time: 95 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 02:54:38 ; Search time 393 Seconds
(without alignments)
9013.724 Million cell updates/sec

Title: US-09-937-060A-19
Perfect score: 1573
Sequence: 1 gccccacagcgcgccctgcg.....gacctgaataaaaaaaaaa 1573

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001D.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1573	100.0	1573	21	AAA75676
2	1547	98.3	1821	22	AAH15652
3	1539.4	97.9	1772	22	AAH16415
4	1539	97.8	1719	22	AAH04477
5	1511.6	96.1	1600	22	AAH14424
6	1479.2	94.0	1533	20	AAH84490
7	1479.2	94.0	1533	22	ABA83273
8	1430.2	90.9	1562	24	ABL90618
9	1429.4	90.9	1447	21	AAA50508

10	1428.4	90.8	1438	22	AA159336	Human polynucleoti
11	1168.6	74.3	1205	22	AAC84161	Human sphingosine
12	1165	74.1	1173	24	ABL59532	Human sphingosine
13	1163.2	73.9	1200	22	AA161122	Human polynucleoti
14	853.4	54.3	1815	21	AA247167	Mouse sphingosine
15	853.2	54.2	1759	22	AAD14425	Mouse consensus sp
16	806.2	51.3	1559	21	AA247168	Mouse sphingosine
17	510.2	32.4	587	22	AAH10443	Human cDNA clone (
18	387.6	24.6	394	24	ABL66446	Lung cancer relate
19	368.8	23.4	675	22	AAH07214	Human cDNA clone (
20	319.6	20.3	604	22	AAH04943	Human cDNA for sph
21	262.8	16.7	2380	22	AAH14817	Human sphingosine
22	262.8	16.7	2380	24	ABL59533	Human sphingosine
23	262.8	16.7	2422	23	AAH85331	DNA encoding novel
24	246.8	15.7	2698	22	AAH14816	Mouse cDNA for sph
25	240	15.3	240	22	AAH04478	Conserved region b
26	193	12.3	296	22	AAH24477	Human ovarian PCR-
27	193	12.3	296	22	AAH83083	Human ovarian tumo
28	188.2	12.0	2136	22	AAH94172	Human full-length
29	181.6	11.5	801	22	AAH99234	Human protein enco
30	132.6	8.4	539	22	AAH92477	Human cDNA 3'-end
31	116	7.4	1521	22	AAH08644	Human kinase (PRIN
32	112.8	7.2	1394	24	ABO54324	Human ovarian anti
33	112.8	7.2	1578	23	AAH85330	DNA encoding novel
34	112.8	7.2	11096	22	ABH18278	Human nervous syst
35	110.6	7.0	2830	22	ABL02327	Drosophila melanog
36	105.8	6.7	297	24	ABH22070	Human ORFX polynuc
37	91.4	5.8	2020	23	ABL03297	Drosophila melanog
38	91.4	5.8	4020	23	ABL03296	Drosophila melanog
39	83	5.3	5686	23	ABH02326	Drosophila melanog
40	60	3.8	60	24	ABH1630	Ala-Pro-Ala-Pro co
41	57	3.6	426	24	ABL51829	Ala-Pro-Ala-Pro co
42	52.2	3.3	2744	16	AAO98470	MSPI-containing p
43	51.2	3.3	4231	21	AAH76592	Human ORFX ORF2147
44	51.2	3.3	4231	21	AAA50510	Human sphingosine
45	51.2	3.3	4413	24	ABL40828	Human sphingosine

ALIGNMENTS

RESULT 1
AAA75676
ID AAA75676 standard: DNA: 1573 BP.

AC AAA75676;
XX
XX
22-JAN-2001 (first entry)

DE DNA encoding a human regulator of intracellular phosphorylation.

XX Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma;
XX neurological disorder; Parkinson's disease; demyelinating disease;
XX meningitis; developmental disorder; neuromuscular disorder; cancer;
XX myasthenia gravis; cell proliferative disorder; actinic keratosis;
XX arteriosclerosis; atherosclerosis; leukemia; melanoma; bronchitis;
XX autoimmune disorder; inflammatory disorder; Addison's disease;
XX acquired immunodeficiency disease; allergy; diabetes mellitus;
XX rheumatoid arthritis; microbial infection; trauma; ss.

OS Homo sapiens.

FT Key Location/Qualifiers
FT CDS 130..1284
FT /tag= a
FT /product= "regulator of intracellular phosphorylation"

XX W0200055332-A2.

XX 21-SEP-2000.

XX 17-MAR-2000; 2000WO-US07277.

18-MAR-1999; 99US-0125593.
PR 20-MAY-1999; 99US-0135049.
PR 09-JUL-1999; 99US-0143188.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y,
PI Lu DM, Au-Young J;
XX
DR WPI: 2000-602121/57.
DR P-PSDB; ABB18659.
XX
PT Novel human intracellular phosphorylation regulator polypeptides and
PT polynucleotides for diagnosis, prevention and treatment of
PT neurological, cell proliferative and autoimmune/inflammatory disorders
PT
PS Claim 4, Page 89, 96pp; English.
XX
XX The present sequence encodes a human regulator of intracellular
CC phosphorylation (HRIP). HRIP is useful for screening agonists and
CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
CC are useful for treating a disease or condition associated with
CC decreased or increased expression of functional HRIP. Diseases treated
CC or diagnosed include neurological disorders such as stroke, Parkinson's
CC disease, demyelinating diseases, bacterial and viral meningitis and
CC other developmental disorders of the central nervous system.
CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders
CC such as actinic keratosis, melanoma, myeloma and cancer of the adrenal gland,
CC including leukemia, melanoma, myeloma and cancer of the adrenal gland,
CC bladder, bone, bone marrow, liver, lung muscle, ovary, autoimmune/
CC inflammatory disorder such as Addison's disease, acquired
CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
CC rheumatoid arthritis, microbial infection and trauma.
CC
XX
SQ Sequence 1573 BP; 274 A; 480 C; 514 G; 305 T; 0 other:
Query Match 100.0%; Score 1573; DB 21; Length 1573;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCCACAGCGGGCCCTTCGACGCGCCCTGGGACACCGTAAGAGCTGAAGCAGG 60
DB 1 GCCCCACAGCGGGCCCTTCGACGCGCCCTGGGACACCGTAAGAGCTGAAGCAGG 60
QY 61 AGCCGCGCGACAGGCGACCGCCACAGCGCCAGGACCCCTGGACACCGGAGCCGCG 120
DB 61 AGCCGCGCGACAGGCGACCGCCACAGCGCCAGGACCCCTGGACACCGGAGCCGCG 120
QY 121 GTGAGGTTATGATCCAGCGGGGCGCCCGGGGGTCTCCCGGCGCTGCGCCG 180
DB 121 GTGAGGTTATGATCCAGCGGGGCGCCCGGGGGTCTCCCGGCGCTGCGCCG 180
QY 121 GTGAGGTTATGATCCAGCGGGGCGCCCGGGGGTCTCCCGGCGCTGCGCCG 180
DB 121 GTGAGGTTATGATCCAGCGGGGCGCCCGGGGGTCTCCCGGCGCTGCGCCG 180
QY 181 CTGGTCTGCTGAACCGCGCGGGGCGGCAAGGCAAGGCTTGCACTCTTCGAGGAC 240
DB 181 CTGGTCTGCTGAACCGCGCGGGGCGGCAAGGCAAGGCTTGCACTCTTCGAGGAC 240
QY 241 GTGAGGCGCTTTGGCTGAGAGGCTGAATCTCTTACGCTGATGCTACGACGGCGG 300
DB 241 GTGAGGCGCTTTGGCTGAGAGGCTGAATCTCTTACGCTGATGCTACGACGGCGG 300
QY 241 GTGAGGCGCTTTGGCTGAGAGGCTGAATCTCTTACGCTGATGCTACGACGGCGG 300
DB 241 GTGAGGCGCTTTGGCTGAGAGGCTGAATCTCTTACGCTGATGCTACGACGGCGG 300
QY 301 AACCAAGCGCGGAGCTGCTGCGGCTCGGAGAGCTGGCGCTGGAGCGCTGCTGCTG 360
DB 301 AACCAAGCGCGGAGCTGCTGCGGCTCGGAGAGCTGGCGCTGGAGCGCTGCTGCTG 360
QY 361 ATGCTGAGAGAGGCTGATGCAAGAGTGTGAACGAGGCTCATGAGAGCGGCTGATG 420
DB 361 ATGCTGAGAGAGGCTGATGCAAGAGTGTGAACGAGGCTCATGAGAGCGGCTGATG 420
QY 421 GAGACGCGCATCAGAGAGCGCTGTGAGCTTCCAGAGCGCTGTGGCAAGCGGCTGCA 480
DB 421 GAGACGCGCATCAGAGAGCGCTGTGAGCTTCCAGAGCGCTGTGGCAAGCGGCTGCA 480

QY 481 GGTTCCTTGAACCATTAATGCTGGCTATGAGAGGTCACCATTAAGAGCTCTGACCAAC 540
DB 481 GGTTCCTTGAACCATTAATGCTGGCTATGAGAGGTCACCATTAAGAGCTCTGACCAAC 540
QY 541 TGCAGCTATTGCTGTGCGCGCGGCTGCTGTCAACCATTAAGAGCTCTGCTGACAG 600
DB 541 TGCAGCTATTGCTGTGCGCGCGGCTGCTGTCAACCATTAAGAGCTCTGCTGACAG 600
QY 601 GCTTGGGAGCTGCGCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 GCTTGGGAGCTGCGCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 GACCTAAGAGTGAAGATATCGGCTGCGGAGATGCGCTTCACTGCGGACCTTC 720
DB 661 GACCTAAGAGTGAAGATATCGGCTGCGGAGATGCGCTTCACTGCGGACCTTC 720
QY 721 CTGCTTGGAGCGCTGCGACCTACCTACCGCGGCGAGCTGAGCTCTGCTGAGGAGA 780
DB 721 CTGCTTGGAGCGCTGCGACCTACCTACCGCGGCGAGCTGAGCTCTGCTGAGGAGA 780
QY 781 GTGGGTTCCAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 781 GTGGGTTCCAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 CTGTTGCGACAGAGAGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 841 CTGTTGCGACAGAGAGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 901 GTGTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 901 GTGTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 GCGCGCTGTGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 961 GCGCGCTGTGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1021 ATGCTGCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1021 ATGCTGCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1081 TACTTGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 1081 TACTTGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1141 TTTGAGTGTATGAGGAAATGATGTTAGCGAGGCGCTGCGAGGCGAGTGCACCCAAAC 1200
DB 1141 TTTGAGTGTATGAGGAAATGATGTTAGCGAGGCGCTGCGAGGCGAGTGCACCCAAAC 1200
QY 1201 TACTTGTGATGTCACGCGGTTGCTGAGAGCCCGCGGCTGAGAGCCCGGAGATG 1260
DB 1201 TACTTGTGATGTCACGCGGTTGCTGAGAGCCCGCGGCTGAGAGCCCGGAGATG 1260
QY 1261 CCAGCGCGAGAGAGGCTTATGAGCGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1261 CCAGCGCGAGAGAGGCTTATGAGCGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1321 GGAACCTTCTCTTCTTCTTCTGAGGCTGAGAGGCTTCCACAGCTCTGCTGAGG 1380
DB 1321 GGAACCTTCTCTTCTTCTTCTGAGGCTGAGAGGCTTCCACAGCTCTGCTGAGG 1380
QY 1381 AGACTCTCTGAGAGAGGTTGAGAGGTTGAGAGGTTATGCTTGTGGGAGCAGGCGCAAT 1440
DB 1381 AGACTCTCTGAGAGAGGTTGAGAGGTTGAGAGGTTATGCTTGTGGGAGCAGGCGCAAT 1440
QY 1441 GAAAGTCTGAGGCTCAGAGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 1441 GAAAGTCTGAGGCTCAGAGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
QY 1501 TTGTTCTGAGAGCGCGCCACCGAGAGCAAAATCCAAATTAAGTATGCTCCAGGCTGAA 1560
DB 1501 TTGTTCTGAGAGCGCGCCACCGAGAGCAAAATCCAAATTAAGTATGCTCCAGGCTGAA 1560
QY 1561 AAAAAAAAAAAAAA 1573

Db 1561 AAAAAAAAAA 1573

RESULT 2

AAH15652

ID AAH15652 standard; cDNA; 1821 BP.

AAH15652;

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:13996.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI: 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

Claim 8: SEQ ID 13996; 2537bp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any special methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95883 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 1821 BP; 301 A; 569 C; 614 G; 337 T; 0 other;

Query Match 98.3%; Score 1547; DB 22; Length 1821;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Gaps 1;

Matches 1558; Conservative 0; Indels 1;

QY 1 GCCCCACAGCCGCGCTGCGAGCCCGCCTGGGACAGCCGATTAAGAGCTTAAGGACAG 60
| | | | |
Db 263 GCCCCACAGCCGCGCTGCGAGCCCGCCTGGGACAGCCGATTAAGAGCTTAAGGACAG 322
| | | | |
QY 61 AGCCCCCGCCAGCGGCGAGGG - CCCCACAGCGCCAGGAGACCCCTGGCAGCGGAGCGCG 119
| | | | |
Db 323 AGCCCGCGCCAGCGGCGAGGGCCGCCACAGCGCGAGGAGACCCCTGGCAGCGGAGCGCG 382
| | | | |
QY 120 GGTGAGGTTATGATATCCAGCGGCGGCGCCCGGGGCGTGTCCCGCGCCCTGGCGCGCT 179
| | | | |
Db 383 GGTGAGGTTATGATATCCAGCGGCGGCGCCCGGGGCGTGTCTCCCGCGCCCTGGCGCGCT 442
| | | | |
QY 180 GCTGTGCTGCTGAACCCGCGGCGGCGGCAAGGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 239
| | | | |
Db 443 GCTGTGCTGCTGAACCCGCGGCGGCGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 502
| | | | |
QY 240 CGTGCAGCGCCCTTTTGGCTGAGAGGCTGAATCTCTTACGCTGATGCTCTACGAGCGCGG 299
| | | | |
Db 503 CGTGCAGCGCCCTTTTGGCTGAGAGGCTGAATCTCTTACGCTGATGCTCTACGAGCGCGG 562
| | | | |
QY 300 GAACCAAGCGCGGAGCTGTGTGCGGCTGAGAGAGCTGGGCGCTGGAGAGCTCTGTGTGT 359
| | | | |
Db 563 GAACCAAGCGCGGAGCTGTGTGCGGCTGAGAGAGCTGGGCGCTGGAGAGCTCTGTGTGT 622
| | | | |
QY 360 CATGCTTGAGAGACGGGCTGATGACAGAGGTGTGAACGGGCTCATGAGAGCGGCTGACTG 419
| | | | |
Db 623 CATGCTTGAGAGACGGGCTGATGACAGAGGTGTGAACGGGCTCATGAGAGCGGCTGACTG 682
| | | | |
QY 420 GGNAGACCGCCATTCAGAGACCCCTGTGTAGCTCCAGAGAGCTCTGGCAGCGCTGGC 479
| | | | |
Db 683 GGNAGACCGCCATTCAGAGACCCCTGTGTAGCTCCAGAGAGCTCTGGCAGCGCTGGC 742
| | | | |
QY 480 AGCTTCCTTGAACCAATTAATGCTGGGCTGATGAGAGAGGCTGACCAATGAAGAGCTCTGACCA 539
| | | | |
Db 743 AGCTTCCTTGAACCAATTAATGCTGGGCTGATGAGAGAGGCTGACCAATGAAGAGCTCTGACCA 802
| | | | |
QY 540 CTGCACGATTATGCTGTGCGCGGCTGTGCACCAATGAACCTGCTGTCTGTGCAC 599
| | | | |
Db 803 CTGCACGATTATGCTGTGCGCGGCTGTGCACCAATGAACCTGCTGTCTGTGCAC 862
| | | | |
QY 600 GCGCTTGGGCGTGGCGCTCTCTGTGTGTCAAGCTGTGGGCTGTCAATGCTGATGT 659
| | | | |
Db 863 GCGCTTGGGCGTGGCGCTCTCTGTGTGTCAAGCTGTGGGCTGTCAATGCTGATGT 922
| | | | |
QY 660 GGAACCTAGAGAGTGAAGATATCGGCTGTGGGAGAGTGGCTTACTGTGGGACCTT 719
| | | | |
Db 923 GGAACCTAGAGAGTGAAGATATCGGCTGTGGGAGAGTGGCTTACTGTGGGACCTT 982
| | | | |
QY 720 CCGGCTGTGGGAGCCCTGCGACCTACCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 779
| | | | |
Db 983 CCGGCTGTGGGAGCCCTGCGACCTACCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1042
| | | | |
QY 780 AGTGGTTCAGAGACACCTGTGCTCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 839
| | | | |
Db 1043 AGTGGTTCAGAGACACCTGTGCTCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1102
| | | | |
QY 840 CCTTGTGCGACAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 899
| | | | |
Db 1103 CCTTGTGCGACAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1162
| | | | |
QY 900 TGTGTAGT 959
| | | | |
Db 1163 TGTGTAGT 1222
| | | | |
QY 960 GGGCGGCTGTGACAGT 1019
| | | | |
Db 1223 GGGCGGCTGTGACAGT 1282
| | | | |
QY 1020 CATGCTGTGCGGCTCTCTCTGTGCGCATGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1079
| | | | |
Db 1283 CATGCTGTGCGGCTCTCTCTGTGCGCATGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1342
| | | | |
QY 1080 CTACTTGGTATATGTGCGCGGTGTGCGCTTGTGAGAGCCCAAGAGATGGAGAAAGTGT 1139
| | | | |

Db 1343 CTACTTGTATGATGTCCTGCTGCTTCCCTGAGACCCCAAGGATGGAAAGTGT 1402
 Oy 1140 GTTTCAGTATGATGAGGAAATGATGTTAGGAGCCCTGACAGGCGCAGTCCACCCAAA 1199
 Db 1403 GTTTCAGTATGATGAGGAAATGATGTTAGGAGCCCTGACAGGCGCAGTCCACCCAAA 1462
 Oy 1200 CTACTTGTATGATGATGAGGAGGTTGCTGAGACCCCTGAGTGAAGCCCAACAGAT 1259
 Db 1463 CTACTTGTATGATGATGAGGAGGTTGCTGAGACCCCTGAGTGAAGCCCAACAGAT 1522
 Oy 1260 GCCACCCCGAAGAGCCCTTATGACCCCTGAGGCGGCTGCTTACTTACTTCTTCTG 1319
 Db 1523 GCCACCCCGAAGAGCCCTTATGACCCCTGAGGCGGCTGCTTACTTACTTCTTCTG 1582
 Oy 1320 AGGACCTTCT 1379
 Db 1583 AGGACCTTCT 1642
 Oy 1380 GAGACT 1439
 Db 1643 GAGACT 1702
 Oy 1440 TGAAGTCT 1499
 Db 1703 TGAAGTCT 1762
 Oy 1500 TTTGTTCT 1558
 Db 1763 TTTGTTCT 1821

RESULT 3
 AAH16415
 ID AAH16415 standard; cDNA: 1772 BP.
 AC AAH16415:
 DT 26-JUN-2001 (first entry)
 XX Human cDNA sequence SEQ ID NO:15193.
 DE Human: primer: detection; diagnosis; antisense therapy; gene therapy; ss.
 XX Homo sapiens.
 OS EPI074617-A2.
 PN 07-FEB-2001.
 PD 28-JUL-2000; 2000EP-0116126.
 PF 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI: 2001-318749/34.
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX Claim 8: SEQ ID 15393; 2537pp + CD ROM; English.
 XX The present invention describes primer sets for synthesizing 5602

full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC of an oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

Query Match 97.9%; Score 1539.4; DB 22; Length 1772;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1551; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Oy 1 GCCCCACAGCCCGCCCTGCGAGCCCGCTGGCGACACGATTAAGGAGCTGAAGCCAGG 60
 Db 220 GCCCCACAGCCCGCCCTGCGAGCCCGCTGGCGACACGATTAAGGAGCTGAAGCCAGG 279
 Oy 61 AGCCGCGCGCAGCGGCGACAGC-CGCCACAGCGCCAGCGAGCCCTGCGAGCGGAGCCGCG 119
 Db 280 AGCCGCGCGCAGCGGCGACAGCGCCAGCGAGCCCTGCGAGCGGAGCGGAGCCGCG 339
 Oy 120 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 179
 Db 340 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 399
 Oy 180 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 239
 Db 400 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 459
 Oy 240 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 299
 Db 460 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 519
 Oy 300 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 359
 Db 520 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 579
 Oy 360 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 419
 Db 580 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 639
 Oy 420 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 479
 Db 640 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 699
 Oy 480 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 539
 Db 700 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 759
 Oy 540 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 599
 Db 760 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 819
 Oy 600 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 659
 Db 820 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 879
 Oy 660 GGTGAGAGTTATGATTCACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 719

```

|||||
Db      880  GGACCTAGAGAGTAGAGATCGGCGCTGGGGAGATGCGCTTCACTCTGGGACCTT
          939
Oy      720  CCTGCGTGGCAGCCCTGCGACCTACCGGGCCGACTGGCTTACCTCCCTGTAGAG
          779
Db      940  CCTGCGTGGCAGCCCTGCGACCTACCGGGCCGACTGGCTTACCTCCCTGTAGAG
          999
Oy      780  AGTGGTTCACAAACACCTGCGCTCCCGCTTGTGTCCAGAGAGCCCGGTAGATGACA
          839
Db      1000 AGTGGGTTCCAAAGACACCTGCGCTCCCGCTTGTGTCCAGAGAGCCCGGTAGATGACA
          1059
Oy      840  CCTGTGCGCACTGGAGAGACCACTGCTCTCACTGACAGTGTGCCCCAGACGACTT
          899
Db      1060 CCTGTGCGCACTGGAGAGACCACTGCTCTCACTGACAGTGTGCCCCAGACGACTT
          1119
Oy      900  TGTGCTACTCCCTGGAGACGTCGCACTGCGACCTGGGCGACAGATGTTTCTGCAACCAT
          959
Db      1120 TGTGCTACTCCCTGGAGACGTCGCACTGCGACCTGGGCGAGTGTGCTGCAACCAT
          1179
Oy      960  GGGCGCTGTGCAGCTGGGCGTCATCTGTCTACGTCGCGCGGAGTGTCTGTC
          1019
Db      1180 GGGCGCGCTGTGCAGCTGGGCGTCATCTGTCTACGTCGCGCGGAGTGTCTGTC
          1239
Oy      1020 CATGCTGCTGCGCTCTTCTCTGCGCANTGAGAGAGGCGACATATGAGTATGACCC
          1079
Db      1240 CATGCTGCTGCGCTCTTCTCTGCGCANTGAGAGAGGCGACATATGAGTATGACCC
          1299
Oy      1080 CTACTTGTATATGCGCCGTCGTGCGCTTCCGCTTGGAGCCCAAGATGGAAAGGT
          1139
Db      1300 CTACTTGTATATGCGCCGTCGTGCGCTTCCGCTTGGAGCCCAAGATGGAAAGGT
          1359
Oy      1140 GTTTCAGATGATGGGGAATTGATGTTAGCGAGGCGCTGCAAGGCGCAGTACCCAAA
          1199
Db      1360 GTTTCAGATGATGGGGAATTGATGTTAGCGAGGCGCTGCAAGGCGCAGTACCCAAA
          1419
Oy      1200 CTACTTGTGATGTCACGCGTGGGTGGAGCCCGCCAGCTGGAAGCCCGACAGAT
          1259
Db      1420 CTACTTGTGATGTCACGCGTGGGTGGAGCCCGCCAGCTGGAAGCCCGACAGAT
          1479
Oy      1260 GCCACCGCCCAAGAGCCCTTATGACCCCTGGCGCGCTGTGCTTACTTCTTGC
          1319
Db      1480 GCCACCGCCCAAGAGCCCTTATGACCCCTGGCGCGCTGTGCTTACTTCTTGC
          1539
Oy      1320 AGGACCTTCTCCCTTCCCTAGGGGCTGTCACACAGCTCCGTGGGGGTGGAG
          1379
Db      1540 AGGACCTTCTCCCTTCCCTAGGGGCTGTCACACAGCTCCGTGGGGGTGGAG
          1599
Oy      1380 GAGACTCTCTGAGAGAGGGTGAAGGTGAGGCTATGCTTGGGGGACAGGCCAGAA
          1439
Db      1600 GAGACTCTCTGAGAGAGGGTGAAGGTGAGGCTATGCTTGGGGGACAGGCCAGAA
          1659
Oy      1440 TGAAGTCTCTGGGTGAGAGAGCCAGCTGGCTGGCGCCAGCTGCTTGAAGGCTTCTAG
          1499
Db      1660 TGAAGTCTCTGGGTGAGAGAGCCAGCTGGCTGGCGCCAGCTGCTTGAAGGCTTCTAG
          1719
Oy      1500 TTTGTCTGAGAGCCCGCCACCGCAGACCAATCCAAATAAAGTACATTCC
          1552
Db      1720 TTTGTCTGAGAGCCCGCCACCGCAGACCAATCCAAATAAAGTACATTCC
          1772

```

RESULT 4
AAD04477 standard; cDNA; 1719 BP.

AC AAD04477;

XX 04-JUL-2001 (first entry)

XX Human sphingosine kinase type 1 (hsk1) cDNA.

XX Human: sphingosine kinase type 1; sk1; chromosome 17q25.2;

KW sphingosine-1-phosphate; SIP; drug screening; therapy; haemostasis;

KW thrombosis; allergic reaction; proliferative disease; cancer;

```

KW haematopoietic disorder; leukaemia; cardiovascular disease; stroke;
KW atherosclerosis; coronary artery disease; dyslipidaemia; diabetes;
KW autoimmune disease; inflammatory disease; multiple sclerosis;
KW T helper-1 related disease; chronic obstructive pulmonary disease;
KW asthma; myocardial infarction; neurodegenerative disorder;
KW wound healing; embryogenesis; anticoagulant; cerebroprotective;
KW neuroprotective; antiproliferative; antiarthritic; cyostatic; cardiac;
KW vulnary; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 270..1424
XX FT /tag= a
XX FT /product= "Human sphingosine kinase type 1 (hsk1)"
XX FT /note= "CDS is specifically claimed in claim 2 and
XX FT shown as SEQ ID NO 2"
XX FT polyA_signal
XX FT 1675..1681
XX FT /tag= b
XX FT misc_feature
XX FT 264..273
XX FT /tag= c
XX FT /note= "Translational initiator AUG is in a partial
XX FT kozak consensus"
XX
XX MO200131029-A2.
XX
XX PD 03-MAY-2001.
XX
XX PF 27-OCT-2000; 2000WO-EP09498.
XX
XX PR 28-OCT-1999; 99US-0162307.
XX PR 07-FEB-2000; 2000US-0180525.
XX
XX PA (WARN ) WARNER LAMBERT CO.
XX
XX PI Allen J, Gosink M, Melendez AJ, Takacs L;
XX
XX DR WPI: 2001-300510/31.
XX DR P-PSDB: AAE00924.
XX
XX PT New human sphingosine kinase type 1 gene for screening drug candidates
XX PT particularly inhibitors used for preventing or treating e.g.
XX PT atherosclerosis, thrombosis, asthma and diabetes
XX
XX PS Claim 2: Fig 1; 91pp: English.
XX
XX CC The present sequence is human sphingosine kinase type 1 (hsk1) cDNA.
XX CC The hsk1 gene is located on chromosome 17q25.2. The sk1 converts the
XX CC substrate sphingosine to sphingosine-1-phosphate (SIP). The sk1 gene
XX CC and encoded polypeptide are applicable in screening drug candidates
XX CC particularly inhibitors for preventing or treating disorders such as
XX CC haemostasis, thrombosis, allergic reactions, proliferative diseases
XX CC including cancer, haematopoietic disorders such as leukaemia,
XX CC cardiovascular diseases such as stroke, atherosclerosis and coronary
XX CC artery disease, dyslipidaemia, diabetes including type I and type II
XX CC diabetes, autoimmune and inflammatory diseases such as multiple
XX CC sclerosis, T helper-1 related diseases, chronic obstructive pulmonary
XX CC disease, asthma, myocardial infarction, neurodegenerative disorders,
XX CC natural wound healing processes and embryogenesis.
XX
XX SQ Sequence 1719 BP; 305 A; 529 C; 556 G; 329 T; 0 other:

```

Query Match 97.8%; Score 1539; DB 22; Length 1719;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1548; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

Oy      11  CGGCCCTGGAGCGCCCGCTGGCAGACCGATTAAGAGCTGAAGCAGAGACCCCGCC
          70
Db      151  CGGCCCTGGAGCGCCCGCTGGCAGACCGATTAAGAGCTGAAGCAGAGACCCCGCC
          210
Oy      71  ACGGCGAGCGCCCGCAGCGCCCGCTGGCAGCGGAGCGCGGCTGAGGTTA
          130
Db      211  ACGGCGAGCGCCCGCAGCGCGCGCTGGCAGCGGAGCGCGGCTGAGGTTA
          270

```


XX The present invention relates to sphingosine kinase (Sphk) polypeptides
CC and nucleic acids encoding them. Sphk is useful for treating a Sphk-
CC associated disorder especially cancers such as leukaemia, lymphoma,
CC ovarian, breast, lung, colon, testicular, stomach and skin,
CC atherosclerosis, restenosis or ischemia and cell proliferative disease
CC or disorder associated with vascular diseases. Sphk gene is used in gene
CC therapy and antisense-therapy. Sphingolipids serving as signalling
CC molecules, have recently emerged as regulators of cell growth,
CC differentiation, diverse cell phenotypes and cell death. Activation of
CC Sphk by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
CC endothelial cells. The present sequence is human sphingosine kinase
CC (Sphk) cDNA.

XX Sequence 1600 BP; 265 A; 492 C; 531 G; 311 T; 1 other:

Query Match 96.1%; Score 1511.6; DB 22; Length 1600;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1534; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

QY 1 GCCCCACACCGGCGCTGGAGCGCCGCTGGGACACGATAGAGCTGAAGCAG 60
DB 46 GCCCCACACCGGCGCTGGAGCGCCGCTGGGACACGATAGAGCTGAAGCAG 105
QY 61 AGCGCGCGCCAGCGGCGAGCG-CCCCACAGCGCCAGGAGCCCGCTGGAGCGCGCG 119
DB 106 AGCGCGCGCCAGCGGCGAGCGCGCCCGCCAGCGCGCGCGCGCGCGCGCG 165
QY 120 GGTGCGAGTTATGATGATCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 166 GGTGCGAGTTATGATGATCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 225
QY 180 GGTGCGAGTTATGATGATCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
DB 226 GGTGCGAGTTATGATGATCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 285
QY 240 CGTGACGCGCGCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 299
DB 286 CGTGACGCGCGCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 345
QY 300 GAACACG 359
DB 346 GAACACG 405
QY 360 CATGTGTGAGAGCGGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 419
DB 406 CATGTGTGAGAGCGGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 465
QY 420 GGAGACCG 479
DB 466 GGAGACCG 525
QY 480 AGCTTCTTGAACCATTTATGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAG 539
DB 526 AGCTTCTTGAACCATTTATGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAG 585
QY 540 CTGACGCGATTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 599
DB 586 CTGACGCGATTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 645
QY 600 GGGTTGGGGGCTGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
DB 646 GGGTTGGGGGCTGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705
QY 660 GGAACCTAAGAGTGAAGATGATGCGGCTGCGGAGATGCGGCTTACCTGCGACCT 719
DB 706 GGAACCTAAGAGTGAAGATGATGCGGCTGCGGAGATGCGGCTTACCTGCGACCT 765
QY 720 CCGGCTTGGGCG 779
DB 766 CCGGCTTGGGCG 825
QY 780 AGTGGTTCACAGACCTGCGCTCCCGCTTGTGTGTCACAGCGCGCGCGGTAGATGACA 839

DB 826 AGTGGTTCACAGACCTGCGCTCCCGCGTGTGTGTCACAGCGCGCGCGGTAGATGACA 885
QY 840 CCTTGGGCG 899
DB 886 CCGTGGCG 945
QY 900 TGTGCTAGTCTGCG 959
DB 946 TGTGCTAGTCTGCG 1005
QY 960 GGGCG 1019
DB 1006 GGGCG 1065
QY 1020 CATGCTGCTGCG 1079
DB 1066 CATGCTGCTGCG 1125
QY 1080 CTACTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1139
DB 1126 CTACTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1185
QY 1140 GTTGGAGTGGATGAGGAGATGATGATGATGATGATGATGATGATGATGATGAT 1199
DB 1186 GTTGGAGTGGATGAGGAGATGATGATGATGATGATGATGATGATGATGATGAT 1245
QY 1200 CTACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
DB 1246 CTACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1305
QY 1260 GCCACCG 1319
DB 1306 GCCACCG 1365
QY 1320 AGGACCG 1379
DB 1366 AGGACCG 1425
QY 1380 GAGACCTCTTGGAGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGAT 1439
DB 1426 GAGACCTCTTGGAGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGAT 1485
QY 1440 TGAAGTCTGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1499
DB 1486 TGAAGTCTGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1545
QY 1500 TTTGTTCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1553
DB 1546 TTTGTTCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1599

RESULT 6
AAV84490
ID AAV84490 standard; DNA: 1533 BP.
XX
AC AAV84490:
XX
DT 01-MAR-1999 (first entry)
XX
DE Human secreted protein gene 80 clone HNF454.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
XX developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX cognitive disorder; schizophrenia; prostatic; obesity; osteoclast; thymus;
XX osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX

Db	248	ACGGGGGGGAACTGCTGGTCCGGAGGACTGGGCCCGCTGGAGACCCTTGCTGGTCACTG	307
QY	365	CTGAGACAGGGCTGATGACACAGAGTGGTGAACGGGC-TCATGAGACGGCCTGACTGGGAG	423
Db	308	YTGAGACGGGCTGTATGACAGAGAGTGGTGAACGGGGCTTCATGAGACGGCCTGACTGGGAG	367
QY	424	ACCGCCATCCAGAAGCCCGCTGTACCTTCACAGACAGGCTCTGGCAACGCCCTGGCACT	483
Db	368	ACCGCCATCCAGAAGCCCGCTGTACCTTCACAGACAGGCTCTGGCAACGCCCTGGCACT	427
QY	484	TCCTTGAAACATTATGCTGGCTATGAGCGAGGTCACCAATGAAGACCTCTGAGCAACCTGC	543
Db	428	TCCTTGAAACATTATGCTGGCTATGAGCGAGGTCACCAATGAAGACCTCTGAGCAACCTGC	487
QY	544	ACGGTATTGCTGTGCGCGCCGGCTGCTGATACCCATGAACCTGCTCTGACACAGGCT	603
Db	488	ACGGTATTGCTGTGCGCGCGCGCTGCTGATACCCATGAACCTGCTCTGACACAGGCT	547
QY	604	TCGGGGCTGCGCCTCTTCTGTGTGCTACGCTTGCGCTGGGGCTTCAATTTGATGTGAGAC	663
Db	548	TCGGGGCTGCGCCTCTTCTGTGTGCTACGCTTGCGCTGGGGCTTCAATTTGATGTGAGAC	607
QY	664	CTAAGAGTGAAGATATGCGGCTCTGGGGGAGATGCGCTTACCTCTGGGCACCTTCTG	723
Db	608	CTAAGAGTGAAGATATGCGGCTCTGGGGGAGATGCGCTTACCTCTGGGCACCTTCTG	667
QY	724	CGTGTGACAGCCCTGGGCACACTACCGCGGCCCACTGAGGCTACCTCCCTGTAGGAAGCTG	783
Db	668	CGTGTGACAGCCCTGGGCACACTACCGCGGCCCACTGAGGCTACCTCCCTGTAGGAAGCTG	727
QY	784	GGTTCCAGACACCTGCTCCGCCCTTTGTGTGTACAGACAGGCGCGGTAGATGACACCTT	843
Db	728	GGTTCCAGACACCTGCTCCGCCCTTTGTGTGTGTACAGACAGGCGCGGTAGATGACACCTT	787
QY	844	GTGGCACATGAGAGAGCCCATGAGGCCCTCTCACTGTGACAGTGTGCTCCGACAGAGACTTGTG	903
Db	788	GTGGCACATGAGAGAGCCCATGAGGCCCTCTCACTGTGACAGTGTGCTCCGACAGAGACTTGTG	847
QY	904	CTAGTCCCTGGACAGCTGCACTGCACTGCACTGAGGAGTGAATTTTGTGTGACCCCTTGGGC	963
Db	848	CTAGTCCCTGGACAGCTGCACTGCACTGCACTGAGGAGTGAATTTTGTGTGACCCCTTGGGC	907
QY	964	CGGTGTGCACTGCGCTGATGATCTGTTCTTACGTGTGCGGGCGGAGTGTCTGTGTCCATG	1023
Db	908	CGGTGTGCACTGCGCTGATGATCTGTTCTTACGTGTGCGGGCGGAGTGTCTGTGTCCATG	967
QY	1024	CTGCTGGGCTCTTTCCTGTGGCCATGAGAGAGGACAGCATATGAGATGATGCCCTTAC	1083
Db	968	CTGCTGGGCTCTTTCCTGTGGCCATGAGAGAGGACAGCATATGAGATGATGCCCTTAC	1027
QY	1084	TTTGGATATGTGGCCCGGTGTCGGCTTCCGCTTGTGAGACCCCAAGATGTGGAAGGTGTGTT	1143
Db	1028	TTTGGATATGTGGCCCGGTGTCGGCTTCCGCTTGTGAGACCCCAAGATGTGGAAGGTGTGTT	1087
QY	1144	GCAGTGGATGGGGAATTTGATGTTTACGAGAGCCGCTGACAGGCGACAGTGTGACCCAACTAC	1203
Db	1088	GCAGTGGATGGGGAATTTGATGTTTACGAGAGCCGCTGACAGGCGACAGTGTGACCCAACTAC	1147
QY	1204	TTTCTGATGGTCAAGCGTTGCTGTGAGAGCCCGCGCCAGCTGGAAGCCCAAGCATGTGCCA	1263
Db	1148	TTTCTGATGGTCAAGCGTTGCTGTGAGAGCCCGCGCCAGCTGGAAGCCCAAGCATGTGCCA	1207
QY	1264	CCGCGAGAGAGCCCTTATAGACCCGCTGGGCGCGGTGTGCTCTTATGTGTCTACTGTCAAGA	1323
Db	1208	CCGCGAGAGAGCCCTTATAGACCCGCTGGGCGCGGTGTGCTCTTATGTGTCTACTGTCAAGA	1267
QY	1324	CCCTTCCCTCTTCCATGAGGCTGACAGGCTGTGTCAACACCTCTGTGGGGGTGAGAGAGA	1383
Db	1268	CCCTTCCCTCTTCCATGAGGCTGACAGGCTGTGTCAACACCTCTGTGGGGGTGAGAGAGA	1327
QY	1384	CTTCTGTGAGAGAGGTGAGAGGTGAGGCTATGTGCTTTGGGGGACAGGCCAGATGAA	1443

Db	1328	CTCCTCTGAGAAAGGTTGAGAAAGTGGAGGCTTTCGTTTGGGGGACAGGCCAGAAATGAA	1387
Qy	1444	GTCTTGCGTACAGAGCCCAAGCTGGCTGGGCCAGCTCCATGTAAAGCCCTTCAGTTTG	1503
Db	1388	GTCTTGCGTACAGAGAGCCAGAGCTGGCTGGGCCAGCTGCGCTATGTAAAGCCCTTCAGTTTG	1447
Qy	1504	TTCCTGAGACCCCGCCACCCAGAAACCAATTCAAATAAGTGACATTTCCAGCCTGAAAAA	1563
Db	1448	TTCCTGAGACCCCGCCACCCAGAAACCAATTCAAATAAGTGACATTTCCCAAAAAAAAAA	1507
Qy	1564	AAAAAAAAAAAA 1573	
Db	1508	AAAAAAAAAAAA 1517	
RESULT 7			
ID	ABA83273	standard; cDNA; 1533 BP.	
AC	ABA83273;		
XX	07-FEB-2002	(first entry)	
DT			
XX			
DE		Human secreted protein gene 80 SEQ ID NO:90.	
XX			
KW		Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;	
KW		dermatological; immunosuppressive; antiinflammatory; immunostimulant;	
KW		cytostatic; cardiant; vascular; anti-angiogenic; ophthalmological;	
KW		neuroprotective; neurotropic; anticonvulsant; antialzheimers; vulnerrary;	
KW		antiparkinsonian; antimetabolic; gene therapy; vaccine; immune disorder;	
KW		multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;	
KW		human immunodeficiency virus; hyperproliferative disorder; wound healing;	
KW		Gaucher's disease; cardiovascular disease; Schmitz syndrome; chemotaxis;	
KW		Chaga's cardiomyopathy; coronary arteriosclerosis; angogenic disorder;	
KW		corneal graft neovascularisation; diabetic retinopathy; regeneration;	
KW		neurological disorder; Huntington's chorea; Alzheimer's disease;	
KW		Parkinson's disease; infectious disease; chromosome 17; ss.	
OS		Homo sapiens.	
XX			
XX			
PN	WO200162891-A2.		
XX			
XX	30-AUG-2001.		
XX			
XX	21-FEB-2001; 2001WO-US05614.		
PR	24-FEB-2000; 2000US-184836P.		
XX	29-MAR-2000; 2000US-193170P.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Ni J, Ebner R, Lafleur DW, Moore PA, Olsen HS, Rosen CA;		
PI	Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;		
PI	Florence C, Hu J, Li Y, Kyaw H, Fischer CL, Ferrle AM, Fan P;		
PI	Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G;		
PI	Zeng Z, Greene JM;		
XX			
DR	WPI: 2001-625724/72.		
DR	P-PSDB; ABB50380.		
XX			
PT	Nucleic acids encoding 207 human secreted polypeptides, useful for		
PT	preventing, diagnosing and/or treating, e.g. cancers, Parkinson's		
PT	disease and diabetic retinopathy -		
XX			
PS	Claim 1; Page 965; 1533pp; English.		
XX			
CC	ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted		
CC	proteins (I) and polynucleotide (II) sequences. (I) and (II) have various		
CC	activities based on the tissues and cells the genes are expressed in.		
CC	Examples of these activities include: immunomodulatory; antisclerotic;		
CC	dermatological; immunosuppressive; antiinflammatory; immunostimulant;		
CC	anti-HIV; cytostatic; cardiant; anti-angiogenic; ophthalmological;		
CC	neuroprotective; neurotropic; anticonvulsant; antialzheimers; vascular;		

XX 05-DEC-2000 (first entry)
DT
DE Human sphingosine kinase A cDNA.
XX
XX Sphingosine kinase A; SKA; human; drug screening; infection;
XX anti-inflammatory; anti-allergic; anticancer; inflammation; allergy;
XX cancer; therapy; diagnosis; ds.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 7..1161
XX /*tag= a
XX
XX MO200052173-A2.
XX
XX 08-SEP-2000.
XX
XX 02-MAR-2000; 2000MO-CA00223.
XX
XX 02-MAR-1999; 99US-0122516.
XX (ALX) NPS ALLELIX CORP.
XX
XX Munroe D, Gupta A, Falzone GR;
XX WPI: 2000-572185/53.
XX P-PSDB: AAY06057.
XX
XX New human sphingosine kinase A, B and C polynucleotides and
XX polypeptides useful in e.g. chromosome and gene mapping, and detecting
XX inflammation or disease associated with abnormal levels of sphingosine
XX kinase expression -
XX
XX Disclosure: Fig 1; 81pp; English.
XX
XX The present sequence is that of an isolated polynucleotide encoding
XX human sphingosine kinase A (SKA, see AAY06057), an enzyme that
XX phosphorylates sphingosine to form sphingosine 1-phosphate.
XX The polynucleotide was isolated from an HeLa cDNA library by
XX PCR amplification. The invention provides polynucleotides (see
XX AAY06057-10) and polypeptides (see AAY06057-53) for the human
XX sphingosine kinase (SK) homologues SKA, SKB and SKC. The
XX polynucleotides may be used as hybridization probes, in the
XX construction of PCR primers for chromosome and gene mapping, in
XX the recombinant production of SKA, SKB and SKC, and in the
XX generation of antisense DNA or RNA. They can be used to detect
XX inflammation or disease associated with abnormal levels of SK
XX expression, or to detect differences in gene sequence between
XX normal and carrier or affected individuals. Host cells expressing
XX SK can be used in drug screening. Human SK specific antibodies,
XX inhibitors, ligands or their analogues can be used as bioactive
XX agents to treat inflammation or disease including viral, bacterial
XX or fungal infections, allergic responses, mechanical injury
XX associated with trauma, hereditary diseases, lymphoma or carcinoma,
XX and other conditions with activate the genes of kidney, lung,
XX heart, lymphoid or tissues of the nervous system.
XX
XX Sequence 1447 BP: 256 A; 427 C; 465 G; 299 T; 0 other;
SQ
Query Match 90.94; Score 1429.4; DB 21; Length 1447;
Best Local Similarity 99.28; Pred. No. 6; 3e-310;
Matches 1436; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 124 GAGGTATGATCGACGGGGGCGGCGGCTCTCCGGGCGCTGCGCGGCTG 183
DB 1 GAGGTATGATCGACGGGGGCGGCGGCGGCTCTCCGGGCGCTGCGCGGCTG 60
QY 184 GTGCTCTGAACCCCGCGCGCGCAAGGCGCAAGGCTTTCACACTTTCGAGTCACTG 243
DB 61 GTGCTCTGAACCCCGCGCGCGCAAGGCGCAAGGCTTTCACACTTTCGAGTCACTG 120

QY 244 CAGCCCTTTTGGCTGAGGCTGAATCTCTTACACCTGATGCTCACTGAGCGCGGAC 303
DB 121 CAGCCCTTTTGGCTGAGGCTGAATCTCTTACACCTGATGCTCACTGAGCGCGGAC 180
QY 304 CAGCGCGGAGCTGTGCGGTTCGAGAGAGCTGGCGCTTGGACCTCTGTGTGCTATG 363
DB 181 CAGCGCGGAGCTGTGCGGTTCGAGAGAGCTGGCGCTTGGACCTCTGTGTGCTATG 240
QY 364 TGTGAGACGGGCTGATGACAGAGCTGTGAACGGCTCATGTGAGCGGCTGACTGAGAG 423
DB 241 TGTGAGACGGGCTGATGACAGAGCTGTGAACGGGCTCATGTGAGCGGCTGACTGAGAG 300
QY 424 ACCGCTATCCAGAAAGCCCTGTGTAGCTCCAGAGAGCTGTGGCAAGCGGCTGAGCT 483
DB 301 ACCGCTATCCAGAAAGCCCTGTGTATCTCTCCAGAGAGCTGTGGCAAGCGGCTGAGCT 360
QY 484 TCTTTGAACGATTTATGCTGTGCTATGAGAGGTCACCAATGAAAGACCTCTGACAACTGC 543
DB 361 TCTTTGAACGATTTATGCTGTGCTATGAGAGGTCACCAATGAAAGACCTCTGACAACTGC 420
QY 544 ACGCTATTGCTGTGCGCGCGCTGCTGTACCAATGAACCTGCTGTCTGTGACACGCT 603
DB 421 ACGCTATTGCTGTGCGCGCGCTGCTGTACCAATGAACCTGCTGTCTGTGACACGCT 480
QY 604 TCGGGGCTGCGCCCTTCTCTGTGCTCAACCCCTGGCGCTTTCATTTGCTGATGTGAC 663
DB 481 TCGGGGCTGCGCCCTTCTCTGTGCTCAACCCCTGGCGCTTTCATTTGCTGATGTGAC 540
QY 664 CTAGAGATGAGAGTATCGGCTGTGGGAGATGCGCTTACCTGTGGGACCTTCTCTG 723
DB 541 CTAGAGATGAGAGTATCGGCTGTGGGAGATGCGCTTACCTGTGGGACCTTCTCTG 600
QY 724 CGTCTGGAGCCCTGGCCACCTTACCGCGGCGGCTTACCTGTGGGAGAGAGT 783
DB 601 CGTCTGGAGCCCTGGCCACCTTACCGCGGCGGCTTACCTGTGGGAGAGAGT 660
QY 784 GGTTCGACAGACCTGCTCCCGGTTGTGTGTCAGAGGCGCGGCTTGTGACACCTT 843
DB 661 GGTTCGACAGACCTGCTCCCGGTTGTGTGTCAGAGGCGCGGCTTGTGACACCTT 720
QY 844 GTGCGACGTGAGAGGCGGCTGCTTCTGCTGACAGAGTGTGCGGACGAGAGCTTGTG 903
DB 721 GTGCGACGTGAGAGGCGGCTGCTTCTGCTGACAGAGTGTGCGGACGAGAGCTTGTG 780
QY 904 CTAGTCTGCGACCTGTGCACTGTGCACTGTGGGAGTGTGCTGACCATGGGC 963
DB 781 CTAGTCTGCGACCTGTGCACTGTGCACTGTGGGAGTGTGCTGACCATGGGC 840
QY 964 CGCTGTGACGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1023
DB 841 CGCTGTGACGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 1024 CTGCTGCGCCCTTCTCTGCGGCAATGAGAGGCGGCAATGATGATGATGATGATGATG 1083
DB 901 CTGCTGCGCCCTTCTCTGCGGCAATGAGAGGCGGCAATGATGATGATGATGATGATG 960
QY 1084 TTGCTATGATGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 1143
DB 961 TTGCTATGATGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 1020
QY 1144 GCAGTGTATGGGGAATGTATGTATGACAGGCGCTGTGACAGGCGGCAATGATGATG 1203
DB 1021 GCAGTGTATGGGGAATGTATGTATGACAGGCGCTGTGACAGGCGGCAATGATGATG 1080
QY 1204 TTCTGATGTGACAGGCTGTGAGAGGCGGCGGCTGTGAGAGGCGGCAATGATGATG 1263
DB 1081 TTCTGATGTGACAGGCTGTGAGAGGCGGCGGCTGTGAGAGGCGGCAATGATGATG 1140
QY 1264 CCGCAGAGAGGCTTATGACAGGCGGCGGCTGTGAGAGGCGGCAATGATGATGATG 1323
DB 1141 CCGCAGAGAGGCTTATGACAGGCGGCGGCTGTGAGAGGCGGCAATGATGATGATG 1200
QY 1324 CCGTTCCT 1383

Dh 667 ACCCTCCCTGTAGGAAGAGTGGGCTCCAGACACCTGCTCCCTCCCTGTTGTGCTCAACAGG 726
Oy 824 GCCCGGTAGATGACACCTTTGTGACCTGAGAGACAGTGCCTCTCTCATCTGAGACGTGG 883
Dh 727 GCCCGGTATATGACACCTTTGTGACCTGAGAGACAGTGCCTCTCTCATCTGAGACGTGG 786
Oy 884 TCCCGGACGAGAGACTTTGTGCTGAGTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 943
Dh 787 TCCCGGACGAGAGACTTTGTGCTGAGTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 846
Oy 944 TCTTTGTCTGACACCTATGAGGCGCTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1003
Dh 847 TCTTTGTCTGACACCTATGAGGCGCTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 906
Oy 1004 CGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063
Dh 907 CGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
Oy 1064 TGGAGTATGAAATGCCCTTACTTGTATATGATGATGATGATGATGATGATGATGATGATGAT 1123
Dh 967 TGGAGTATGAAATGCCCTTACTTGTATATGATGATGATGATGATGATGATGATGATGATGAT 1026
Oy 1124 AGGATGAGGAAGGTGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1183
Dh 1027 AGGATGAGGAAGGTGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1086
Oy 1184 GCGAGTGTGACCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1243
Dh 1087 GCGAGTGTGACCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
Oy 1244 GGAAGCCCGGAGAGATGCCCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1286
Dh 1147 GGAAGCCCGGAGAGATGCCCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1189

RESULT 12

ABL59532 standard; cDNA: 1173 BP.

ABL59532;

16-JUL-2002 (first entry)

Human sphingosine kinase (SphK1) cDNA SEQ ID NO:32.

Human: sphingosine kinase; SphK1; enzyme; chromosome 17q25.2; gene;

tumour; lipid associated gene; lipid metabolism; lipid synthesis; ss.

Homo sapiens.

WO200227028-A1.

04-APR-2002.

27-SEP-2001; 2001WO-US30366.

28-SEP-2000; 2000US-0676052.

(ATRI-) ATRIRGIN TECHNOLOGIES INC.

Skinner MK, Patton JL, Chaudhary J;

WPI; 2002-402054/43.

Identifying tumor characteristics in a tissue sample taken from a

patient, involves determining the copy number or expression level of

genes associated with lipid metabolism, synthesis or action

Example 1: Page 90; 113pp; English.

The present invention describes a method for identifying tumour

characteristics, comprising measuring a copy number or expression level

of at least two genes associated with lipid metabolism, synthesis, or

CC action in cells from a patient tissue sample, and comparing the results
CC with a copy number or expression level of the genes in a normal cell.
CC Also described is an array of nucleic acid polymers immobilised on a
CC solid support, comprising a solid support, at least two different nucleic
CC acid polymers which are each specific for a different gene associated
CC with lipid metabolism, synthesis or action, where each nucleic acid
CC polymer is located at a predetermined position on the solid support, and
CC the array comprises nucleic acid polymers which are specific for less
CC than 100 genes other than the selected genes. The method is useful for
CC determining tumour characteristics in a tissue sample taken from a
CC patient. The present sequence represents a human lipid-associated gene
CC related cDNA sequence, which is used in the exemplification of the
CC present invention.

xx Sequence 1173 BP; 188 A; 349 C; 390 G; 246 T; 0 other;

Query Match 74.1%; Score 1165; DB 24; Length 1173;

Best Local Similarity 99.6%; Pred. No. 7.3e-251;

Matches 1168; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 112 GAGCGCGGGTGGAGTGTATGATCCAGCGGGGCGCCCGGGCGCTGCTCCCGCGGCC 171
Dh 1 GAGCGCGGGTGGAGTGTATGATCCAGCGGGGCGCCCGGGCGCTGCTCCCGCGGCC 60
Oy 172 TCCCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 231
Dh 61 TCCCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Oy 232 CGGAGTGTGACGAGCCCTTTTGGCTGAGGCTGAAATCTCTTCAAGCTGATGCTACT 291
Dh 121 CGGAGTGTGACGAGCCCTTTTGGCTGAGGCTGAAATCTCTTCAAGCTGATGCTACT 180
Oy 292 GAGCGCGGAGAACCGCGGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 351
Dh 181 GAGCGCGGAGAACCGCGGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Oy 352 CTGAGTGTGATGCTGTGAGAACGGGCTGATGACAGAGTGTGAGAACGGGCTGATGAGAGCG 411
Dh 241 CTGAGTGTGATGCTGTGAGAACGGGCTGATGACAGAGTGTGAGAACGGGCTGATGAGAGCG 300
Oy 412 CCTGACTGGAGAACCGCATTCAGAACCCCTGTGTAGCTCTCCAGAGGCTCTGCGAAC 471
Dh 301 CCTGACTGGAGAACCGCATTCAGAACCCCTGTGTAGCTCTCCAGAGGCTCTGCGAAC 360
Oy 472 GCGCTGGAGCTTCTTGAACCATTTGCTGTATGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 531
Dh 361 GCGCTGGAGCTTCTTGAACCATTTGCTGTATGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 420
Oy 532 CTGACCACTGACAGCTATTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591
Dh 421 CTGACCACTGACAGCTATTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Oy 592 CTGACACAGGCTTCGGGGCGCGCTTCTGTGCTCAAGCTGTGGGCTTCATTT 651
Dh 481 CTGACACAGGCTTCGGGGCGCGCTTCTGTGCTCAAGCTGTGGGCTTCATTT 540
Oy 652 GCTGATGTGAGCTTAGAGAGTGAAGTATCGGCGCTGTGGGGAGTGTGGGCTTCACTGTG 711
Dh 541 GCTGATGTGAGCTTAGAGAGTGAAGTATCGGCGCTGTGGGGAGTGTGGGCTTCACTGTG 600
Oy 712 GGCACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 771
Dh 601 GGCACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Oy 772 GTAGAGAGAGTGGTTCAGAGACCTGCTCCCGTGTGTGTGTCAGAGAGGCGCGGTA 831
Dh 661 GTAGAGAGAGTGGTTCAGAGACCTGCTCCCGTGTGTGTGTCAGAGAGGCGCGGTA 720
Oy 832 GATGACACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
Dh 721 GATGACACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Oy 892 GAGGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951


```
OY 1085 TGGTATATGTCGCCCGTGTGCGCTTCCGCTTGAGGCCAAGATGGAAAGTGTGTTG 1144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 TGGTATAGTGGCCCGTGTGCGCTTCCGCTTGAGGCCAAGATGGAAAGTGTGTTG 780
OY 1145 CAGTGGATGGGAATGATGTTAGGAGCGCGCTGCAGGCGCAGGTGCACCCAACTACT 1204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 781 CAGTGGATGGGAATGATGTTAGGAGCGCGCTGCAGGCGCAGGTGCACCCAACTACT 840
OY 1205 TCTGATGTCAGCGGTTCCGTGAGACCCCGCCCAAGCTGGAAAGCCCGCAGATCCGAC 1264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 841 TCTGATGTCAGCGGTTCCGTGAGACCCCGCCCAAGCTGGAAAGCCCGCAGATCCGAC 900
OY 1265 CCGCAGAAAGACCCCTTATGACCCCTGCGCGGCTGTGCTTACTTACTTGCAGAGAC 1324
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 901 CCGCAGAAAGACCCCTTATGACCCCTGCGCGGCTGTGCTTACTTACTTGCAGAGAC 960
OY 1325 CCTTCCTCCTCCCTAGAGGCTGCAGGCGCTGTGCACAGCTCCTGTGGGGGTGAGAGAC 1384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 CCTTCCTCCTCCCTAGAGGCTGCAGGCGCTGTGCACAGCTCCTGTGGGGGTGAGAGAC 1020
OY 1385 T-CCTCTGAGAAAGGTGAGAAAGTGTGAGGCTATGCTTTGGGGGACAGGCCAGATGAA 1443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 TCCCTCTGAGAAAGGTGAGAAAGTGTGAGGCTATGCTTTGGGGGACAGGCCAGATGAA 1080
OY 1444 GTCTGAGGTCGA-GGAGCCGAGCTGGCTGGGCGCAGCTGCTATGTAAAGCCTTCTAGTTT 1502
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 GTCTGAGGTCGA-GGAGCCGAGCTGGCTGGGCGCAGCTGCTATGTAAAGCCTTCTAGTTT 1140
OY 1503 GTTCTGAGACCCCGCAGCCGACAGCCAAATTCAAATTAAGTACATTTCCAGCCTGAAA 1562
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 GTTCTGAGACCCCGCAGCCGACAGCCAAATTCAAATTAAGTACATTTCCCAAAAAA 1200
```

RESULT 14

AAZ47167 standard; DNA; 1815 BP.

AAZ47167:

28-MAR-2000 (first entry)

Mouse sphingosine kinase 1a DNA.

Anticancer; antiproliferative; antiatherosclerotic; stroke; restenosis;
KMN antineurodegeneration; sphingosine kinase; cell proliferation; apoptosis;
KM cancer; diabetic neuropathy; Alzheimer's disease; atherosclerosis; ds.

Mus musculus.

W09961581-A2.

02-DEC-1999.

25-MAY-1999; 99MO-US11521.

26-MAY-1998; 98US-0086657.

11-AUG-1998; 98US-0096049.

(DEAN-) OFFICE DEAN RES & GRADUATE EDUCATION.

Spiegel S:

WPT: 2000-072612/06.

P-PSDB: AA156053.

New sphingosine kinase, used to treat diseases involving abnormal cell proliferation, e.g. cancer -

Example 1; Page 113-114; 116pp: English.

This sequence represents the coding region for the mouse sphingosine kinase 1a. The sequence was isolated by searching a dbEST (expressed

CC sequence tag database) using amino acid sequence results from sequenced
CC tryptic digests of the rat sphingosine kinase. Expression of sphingosine
CC kinase in cells results in formation of sphingosine-1-phosphate, a known
CC second messenger, and confers serum-independent growth. Increases
CC proliferation, and suppresses serum-deprivation or ceramide-induced
CC apoptosis. The sphingosine kinase nucleic acid is used: (a) to increase
CC sphingosine kinase content of cells, specifically for reducing cell death
CC and/or increasing cell proliferation; and (b) to produce transfected
CC cells that are used to screen for agents that inhibit or promote
CC sphingosine kinase activity. Agents that reduce sphingosine kinase
CC activity or expression are used: (i) to reduce cell proliferation,
CC specifically for treating cancer, and (ii) to treat diseases associated
CC with abnormal cell migration or motility, particularly cancer, restenosis
CC or diabetic neuropathy (but also atherosclerosis, stroke and Alzheimer's
CC disease), whereas agents that stimulate sphingosine kinase can be used
CC to treat conditions associated with reduced cell proliferation, e.g.
CC developmental retardation.

SQ Sequence 1815 BP; 382 A; 511 C; 533 G; 389 T; 0 other:

Query Match 54.3%; Score 853.4; DB 21; Length 1815;

Best Local Similarity 73.9%; Pred. No. 3.5e-181;

Matches 1171; Conservative 0; Mismatches 391; Indels 23; Gaps 6;

```
OY 1 GCGCCACAGCGCGCCTCCGACGCGCGCGCTGGGACACCGAATGAGAGTGAAGCAGG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 238 GCGCCACAGCGCGCCTCCGACACCCCTCCTGGGCAACCGAATGAGAGTGAAGCAGG 297
OY 61 AGCGCGCGCGCAGCGGAGCGGCC---ACAGCGCCAGGAGACCCCTGCGGCGGAGGCC 116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 298 AGCGCGCGGTACTCTTAGCAGCGCGCGGAGCAGCGGTGGCCCTGTGCAGCGGAGGCC 357
OY 117 GCGGCTGAGAGTTATGATTCACGCGGCGCGCGCGCGCTGCTCCGCGCGCTGCGG 176
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 358 CCGGACCTGGCTATGGAACA--GAATGCCCTCGAGAGCTCTCCACGCGCATGACG 414
OY 177 CGTGTGCTGCTCTGTAACCCCGCGGCGGAGGCAAGGCAAGCCTTCAGCTCTCCGAG 236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 415 AGTGTGTGCTCTGTAACCCCGAGGTGAGAGGCAAGGCTCTGAGCTCTTCCAGAG 474
OY 237 TCACTGACAGCCCGCTTTGGCTGAGAGCTGAATCTCCTTACAGCTGATGCTACTGAGCG 296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 475 CCGTGTGCAAGCCCTTCCCTGAGAGGAGGACAGATACCTTTAACTGATACTCACGACG 534
OY 297 GCGGAACACAGCCCGGAGAGCTGTGCGTGCAGAGAGCTGGCGCTGGAGCGCTTGGT 356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 535 GAAGAACCATGTCAGGAGAGCTGTGTGTCAGAGAGTTGGTCACTGAGGACCGCTGGC 594
OY 357 GGTCAATGCTGAGAGCGGCGCTGATGACGAGGAGTGAACGGGCTGATGAGGCGCTGA 416
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 595 AGTCATGTCCGGATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 654
OY 417 CTGGGAGACCGCCATCCAGAGCCCGCTGTAGCTCCAGAGAGCTTGGCAACCGCT 476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 655 CTGGGAGAGCTGCGCATCCAGAAACCCCTGTGTAGCTCCCTGGAGAGCTCCGCAATGCGCT 714
OY 477 GCGAGCTTCCCTTGAACATTATGCTGCTAGAGCAGAGTACCACAAATGAAGACCTCTGAC 536
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 715 GCGAGCTTCTGTGAACACATATGCTGGTGAAGAGAGTGAATGAAGACCTGCTCAT 774
OY 537 CAACGTGACAGCTATGCTGTGCGCGCGGCTGCTGTACCCATTAACCTGTGTCTGCA 596
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 775 CAACGTGACAGCTGTGTGTGCGCGCGGCTGCTGTACCCATTAACCTGTGTCTGCA 834
OY 597 CAGGCTTGGGCGCTGCGCTTCTGTGTGCTCACCTGCGGCGCTTCACTTCTGCA 656
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 835 CACTGCTTGGGCTGCGGCTGCTATCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 894
OY 657 TGTGACCTAGAGAGTGAAGATGAGCGCTGTGGGAGAGATGCGCTTCACTTGTGGCAC 716
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 895 CGTGAACCTGAGAGTGAAGATGAGCGCTTGGGAGAGATTCGTTTCACTGTGGGCGAC 954
OY 717 CTTCTGCGCTGTGAGACCGCTGCGACCTAAGCGCGGCGCACTGCGCTTACCTGCTTAGG 776
```

```

Db 955 CTTCTTTGGCCTACCAAGCCTGGCATCTACCAAGCCACTGGCCCTTCTCTAGG 1014
Oy 777 AAGAGGGTTTCCAAAGACCTGCTCCCGTGTGTGTGCAGAGGGCCGCTAATAC 836
Db 1015 AACTGTGGCTCTTAAGAGACCCGCTCTAC---ACTGTGAGAGAGGCCCGCTGCACAC 1071
Oy 837 ACACCTGTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 896
Db 1072 ACACCTGTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1131
Oy 897 CTTTGTGTACTCTCTGACACTGCTGACATGCACTGGAGAGAGAGAGAGAGAGAG 956
Db 1132 CTTTGTGTACTCTCTGACACTGCTGACATGCACTGGAGAGAGAGAGAGAGAGAG 1191
Oy 957 CATTGGGCGCTGTGACACTGCTGACATGCTGTTTACGTGGGCGGAGAGAGAGAG 1016
Db 1192 CATTGGGCGCTGTGACACTGCTGACATGCTGTTTACGTGGGCGGAGAGAGAGAG 1251
Oy 1017 TCCCATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1076
Db 1252 GCGTGGCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1311
Oy 1077 CCCCCTACTGTTATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1136
Db 1312 TCCCATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1371
Oy 1137 TGTGTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1196
Db 1372 CGTGTTTCTGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1431
Oy 1197 AAACCTACTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1256
Db 1432 AAACCTACTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1491
Oy 1257 GATGACACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1313
Db 1492 GGGGCGACCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1551
Oy 1314 ACTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1373
Db 1552 GGGGCGACCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1611
Oy 1374 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1433
Db 1612 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1670
Oy 1434 CCAGATGATGATCTGTGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1484
Db 1671 TCTACCTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1730
Oy 1485 TGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1544
Db 1731 TGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1790
Oy 1545 ACATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1569
Db 1791 ACATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1815

```

```

KW Leukemia: vasotropic; cell proliferative disorder; vascular disease; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT CDS 327..1475
FT /tag="a
FT /product="Mouse sphingosine kinase (Sphk) protein"
FN W0200160990-A2.
XX 23-AUG-2001.
XX 14-FEB-2001; 2001WO-US04789.
XX 14-FEB-2000; 2000US-0182360.
XX 22-MAR-2000; 2000US-0191261.
XX (CURA-) CUPAGEN CORP.
XX (GUTH) GENENTECH INC.
XX Rastelli L.
XX MPI: 2001-514770/56.
XX P-PSDB; AAB07883.
XX An isolated Sphingosine kinase polypeptide useful for treating a
XX Sphk-associated disorder especially cancer, restenosis or ischemia in a
XX human.
XX Example 1; Page 92; 107pp; English.
XX The present invention relates to sphingosine kinase (Sphk) polypeptides
XX and nucleic acids encoding them. Sphk is useful for treating a Sphk-
XX associated disorder especially cancers such as leukemia, lymphoma,
XX ovarian breast, lung polyp, esophageal, stomach and skin.
XX atherosclerosis, restenosis or ischemic diseases. Sphk gene is used in gene
XX therapy and antisense-therapy. Sphingosine kinase is a signaling
XX molecule, have recently emerged as regulators of cell growth. Activation of
XX Sphk by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
XX endothelial cells. The present sequence is consensus mouse sphingosine
XX kinase (Sphk) cDNA.
XX
XX Sequence 1759 BP; 357 A; 495 C; 523 G; 384 T; 0 other;
XX
XX Query Match 54.28; Score 853.2; DB 22; Length 1759;
XX Best Local Similarity 73.98; Pred. No. 3,9e-181;
XX Matches 1158; Conservative 0; Mismatches 388; Indels 22; Gaps 5;
Oy 3 CCCACAGCCGCGCCCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 62
Db 196 CCCACAGCGAGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 255
Oy 63 CCGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 118
Db 256 CCGCGCTTACCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 315
Oy 119 GAGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 178
Db 316 GAGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375
Oy 179 TGTGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 238
Db 376 TGTGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
Oy 239 AGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 298
Db 436 GTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 495
Oy 299 GGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 358

```

RESULT 15
 AADI4425
 ID AADI4425 standard; cDNA: 1759 BP.
 AC AADI4425;
 XX 01-NOV-2001 (first entry)
 XX
 XX Mouse consensus sphingosine kinase (Sphk) cDNA.
 XX
 XX Mouse: sphingosine kinase; Sphk; restenosis; ischemia; gene therapy;
 XX anti-sense therapy; cancer; sphingolipid; signalling molecule; apoptosis;
 XX cytoskeletal; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;
 KM

Db 486 AGAACCAATGCCAGGAGCTGTGTGTGCAGAGAGCTTGGGTCACTGGGAGCGCCCTGGCAG 555
Qy 359 TCATGTCTGAGACGGGCTGATGCACAGAGTGTGTGAACGGGCTCATAGAGCGGCTGACT 418
Db 556 TCATGTCCGGTATGTGTGATGCATGAAGTGTGTGAATGGGCTAATGGAACGGCCAGACT 615
Qy 419 GGGAGACCGCCATCCAGAGAGCCCTGTGTAGCCCTCCACAGAGGCTCTGGCAACGGCCTGG 478
Db 616 GGGAGACGTCATCCAGAGAGCCCTGTGTAGCCCTCCACAGAGGCTCCGGCAATGGCCTGG 675
Qy 479 CAGCTTCCTTGAACATTTATGTGCTATGACAGAGTCAACCAATGAAGCTCTCTGACCA 538
Db 676 CAGCTTCCTTGAACATTTATGTGCTATGACAGAGTCAACCAATGAAGCTCTCTGACCA 735
Qy 539 ACTGACGCTATTTGCTGTGACCGCCGCTGTGCTGACCCCTGAGAACCTGCTGTCTGACCA 598
Db 736 ACTGACACACTGCTGTGTGTGACCGCCGCTGTGACCCCTGAGAACCTGCTGTCTGACCA 795
Qy 599 CGGCTTCGGGGCTGGCCCTCTCTGTGTGCTGACCTGGCCCTGGGGCTTCATTTGTGATG 658
Db 796 CTGCTTCGGGGCTGGGGCTCTATTCGTGTGCTGACGTCTGTCTGGGGCTTTGTTGCTGAG 855
Qy 659 TGGACCTAGAGAGTGAAGATTCGGCCCTGTGGGGAGATGCGCTTCACTCTGGCCACT 718
Db 856 TGGACCTGAGAGTGAAGATTCGGCCCTGTGGGGAGATTCGTTTCACAGTGGGCACCT 915
Qy 719 TCTGTGCTGTGACAGCCCTGCGGCTGACCGGCGGCTGACCTGCTGCTGTAGAA 778
Db 916 TCTGTGCTGTGACAGCCCTGCGGCTGACCGGCGGCTGACCGGCTGCTGCTGTAGAA 975
Qy 779 GAGTGGGTTCCAAGACACCTGCTCCCGCTTGTGTGCTGACAGAGGCGCGGTAGATGAC 838
Db 976 CTGTGGCCCTCTAAGAGAGCCGCGCTTAC - - - ACTGTGACAGAGGCGCGCGTGCACAC 1032
Qy 839 ACCTGTGCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 898
Db 1033 ACCTGTGCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1092
Qy 899 TGTGTGTATGCTGTGACACTGCTGACCTGCGGAGTGAATGTTGCTGACACCA 958
Db 1093 TGTGTGTATGCTGTGACACTGCTGACCTGCGGAGTGAATGTTGCTGACACCA 1152
Qy 959 TGGCCGCTGTGACGCTGCTGATGATCTGTTACGTGCGGCGGAGGAGTGTCTGCTG 1018
Db 1153 TGGCCGCTGTGACGCTGCTGATGATCTGTTACGTGCGGCGGAGTGTGTCAAGGG 1212
Qy 1019 CCATGCTGCTGCGCTCTTCTGCGCATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1078
Db 1213 CTGCGCTGCTGCGCTCTTCTGCGCATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1272
Qy 1079 CCTACTTGTATATGTCGCGGCTGCTGCGCTTGGAGGCCAAGAGATGGAGAAAGTG 1138
Db 1273 CATACCTGCTATGTCGCGGCTGCTGCGCTTGGAGGCCAAGAGATGGAGAAAGTG 1332
Qy 1139 TGTTCAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1198
Db 1333 TGTTCAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1392
Qy 1199 ACTACTTCTGATGCTGACGCTTGTGCGCATGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1258
Db 1393 ACTACTTCTGATGCTGACGCTTGTGCGCATGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1452
Qy 1259 TGGCAACCGCAGAGAGGCT - - - TATGACCCCTGGCGCGGCTGTGCTTGTGCTTAC 1315
Db 1453 GGCACACCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1512
Qy 1316 TTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1375
Db 1513 GAGCCCTCCCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1572
Qy 1376 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1435
Db 1573 GGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1630

Qy 1436 AGAATGAGTCTGTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1485
Db 1631 TTACCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1690
Qy 1486 GTAAGGCTTCTAGTTTGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1545
Db 1691 AGAAGACATTCCTGCTTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1750
Qy 1546 CATTCCTCA 1553
Db 1751 CTTTTCCTCA 1758

Search completed: May 17, 2003, 11:14:20
Job time : 412 secs

•
•
•



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:50:29 ; Search time 14 Seconds
(without alignments)
1137.636 Million cell updates/sec

Title: US-09-937-060A-5

Perfect score: 2016
Sequence: 1 MDPAGPGRGVLPKPCRVLV.....CPEPPSWKPGQMPPEEPL 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2016	100.0	384	1 SPH1_HUMAN	Q9NYA1 homo sapien
2	892	44.2	617	1 SPH2_MOUSE	Q9J1A7 mus musculu
3	886.5	44.0	654	1 SPH2_HUMAN	Q9NPA0 homo sapien
4	103	5.1	297	1 BMRU_BACSU	P39074 bacillus su
5	100.5	5.0	433	1 Y036_STYNY	O55452 synechocyst
6	99.5	4.9	1065	1 KDC1_HUMAN	O75912 homo sapien
7	93.5	4.6	791	1 KDC1_MOUSE	O01383 drosophila
8	93	4.6	712	1 DNLJ_RHOMR	P49421 rhodothermu
9	92.5	4.6	439	1 YSM3_CAEEL	Q10123 caenorhabd
10	92.5	4.6	492	1 GVD2_HALNI	Q9JHT2 halobacteri
11	92	4.6	782	1 SMAB_MOUSE	O62179 mus musculu
12	92	4.6	1473	1 NAL1_HUMAN	O9C000 homo sapien
13	89	4.4	414	1 PRS6_CAEEL	P46502 caenorhabd
14	89	4.4	533	1 SERA_RAT	O08651 rattus norv
15	87	4.3	533	1 SERA_HUMAN	O43175 homo sapien
16	87	4.3	572	1 POXB_ECOLI	P07003 escherichia
17	86.5	4.3	2470	1 NTC2_MOUSE	O35516 mus musculu
18	86	4.3	390	1 ARP2_SCHPO	O9UWJ1 schizosacch
19	86	4.3	448	1 5H7_MOUSE	P32304 mus musculu
20	85.5	4.2	662	1 LOX2_PIG	P16659 sus scrofa
21	85	4.2	608	1 HFAFC_CAUCR	Q45978 caulobacter
22	85	4.2	827	1 KDGL_CAEEL	O03603 caenorhabd
23	84.5	4.2	335	1 G3P2_METAC	P58833 methanosarc
24	84.5	4.2	369	1 XYLM_PSEPU	P21395 pseudomonas
25	84.5	4.2	471	1 HH_DROME	Q02936 drosophila
26	84.5	4.2	492	1 GVD2_HALNI	P33358 halobacteri
27	83.5	4.1	586	1 5H7_RAT	Q9UWJ1 schizosacch
28	83	4.1	547	1 TUI2_SCHPO	P32305 rattus norv
29	83	4.1	520	1 CMCH_MYCIT	O51080 nocardia la
30	83	4.1	1247	1 IRBP_HUMAN	P10745 homo sapien
31	82	4.1	668	1 ACOR_AICER	P28614 alcaligenes
32	82	4.1	873	1 PODK_CLOSY	P22983 clostridium
33	82	4.1	942	1 KDCT_HUMAN	P52824 homo sapien

34	82	4.1	4568	1 DYHB_CHLRE	O39565 chlamydomon
35	81.5	4.0	334	1 G3P_PYRAB	Q9Y1P1 pyrococcus
36	81.5	4.0	392	1 RURE_AICCA	P42454 acinetobact
37	81.5	4.0	554	1 CXIA_PAROE	P08305 paracoccus
38	81	4.0	425	1 BIDA_SERMA	P36568 serratia ma
39	81	4.0	521	1 CMCH_STRCL	O85728 streptomyce
40	81	4.0	771	1 CTPG_MYCIT	O10866 mycobacteri
41	81	4.0	1434	1 VG65_HSVII	Q00106 ictalurid h
42	80.5	4.0	311	1 VGS_ECOLI	P03030 escherichia
43	80.5	4.0	801	1 DHGA_AICCA	P05465 acinetobact
44	80.5	4.0	978	1 PEX6_RAT	P54777 rattus norv
45	80.5	4.0	1047	1 EF3_SCHPO	O94489 schizosacch

ALIGNMENTS

RESULT 1	ID	SPH1_HUMAN	STANDARD:	PRT:	384 AA.
AC	Q9NYA1	Q9NYA13: Q9HPD2: Q9NY70:			
DT	15-JUN-2002	(Rel. 41, Created)			
DT	15-JUN-2002	(Rel. 41, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Sphingosine kinase 1 (EC 2.7.1.-) (SK 1) (SPK 1).				
GN	SPHK1 OR SPHK OR SPK				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20323213; PubMed=10863092;				
RA	Melendez A.J., Carlos-Dias E., Gosink M., Allen J.M., Takacs L.;				
RT	"Human sphingosine kinase: molecular cloning, functional				
RT	characterization and tissue distribution.";				
RL	Gene 251:19-26(2000).				
RN	[2]				
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.				
RX	MEDLINE=20263733; PubMed=10802064;				
RA	Nava V.E., Lacana E., Poulton S., Liu H., Sugitara M., Kono K.;				
RA	Milstien S., Kohana T., Spiegel S.;				
RT	"Functional characterization of human sphingosine kinase-1.";				
RN	FEBS Lett. 473:81-84(2000).				
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.				
RX	MEDLINE=20407120; PubMed=10947957;				
RA	Pitson S.M., D'Andrea R.J., Vandeleur L., Moretti P.A.B., Xia P.;				
RA	Gamble J.R., Vadas M.A., Waltenberg B.W.;				
RT	"Human sphingosine kinase: purification, molecular cloning and				
RT	characterization of the native and recombinant enzymes.";				
RL	Biochem. J. 350:429-441(2000).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Van Veldhoven P.P., Gijssels S.;				
RT	Submitted (Aug-1999) to the EMBL/Genbank/DBJ databases.				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Ovary, and Mammary gland;				
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.;				
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.;				
RA	Wagatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.;				
RA	Takahashi M., Chiba Y., Ishida S., Murekawa K., Ono Y., Takiguchi S.;				
RA	Matenabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saiko K.;				
RA	Yamamoto K., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.;				
RA	Ninomiya K., Iwayanagi T.;				
RT	"NEO human cDNA sequencing project.";				
RT	Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.				
CC	- FUNCTION: Catalyzes the phosphorylation of sphingosine to form				
CC	sphingosine 1-phosphate (SPP), a lipid mediator with both intra-				
CC	and extracellular functions. Also acts on D-erythro-sphingosine				
CC	and to a lesser extent sphinganine, but not other lipids, such as				
CC	D,L-threo-dihydrosphingosine, N,N-dimethylsphingosine.				


```

DR EMBL: AF415214; AAL07500.1; -
DR EMBL: AK004951; BAB23694.1; -
DR EMBL: BC006941; AAH06941.1; -
DR MGI: 1861380; Sphk2.
DR InterPro: IPR001206; DAGKc.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00781; DAGKc; 1.
DR ProDom: PD005043; DAG_kin_cat; 1.
DR SMART: SM00046; DAGKc; 1.
DR Transferrase: Kinase; ATP-binding.
FT CONFLICT 252 252 N -> S (IN REF. 1).
FT CONFLICT 510 510 P -> T (IN REF. 1).
FT CONFLICT 548 548 L -> F (IN REF. 1).
SQ SEQUENCE 617 AA; 65618 MM; 40BE2C2C28B8E26A CRC64;

Query Match 44.2%; Score 892; DB 1; Length 617;
Best Local Similarity 39.9%; Pred. No. 5.7e-70;
Matches 192; Conservative 60; Mismatches 109; Indels 120; Gaps 5;

QY 10 VLPRPCRVLTILNPRGKGKALQIFRSHVOPDLAEISFTLMTERRNHARELYRSEEL 69
DB 141 LLPRKPRLLILVNPRGGGLAMQRCMDHYVPMISAGLSFNLIQTERQNHARELVQGLST 200
QY 70 GRMDALVWMSGDGLMHEVNGMLMRPDMETAIQKPLCSLPAGSGNALAASLNHYAGYEQY 129
DB 201 SEMEGIVVSGDGLYEVLNGLDRPDWEDAVRMPIGVLPCCSGNALAGAVNHGCFEQY 260
QY 130 TNEDLLTCTLLICRLLSPMHLISLHTASGLRFLFSVLSLWAGFLADVDLSEKXRRLLGE 189
DB 261 VGVDDLKCSLLHCRGGSHPRDLDSVTLASGRCSFSLVWAGFLSDVDHSEFRALGS 320
QY 190 MKFTLGFTLRALAKTYRGRLAYLPV----- 215
DB 321 ARFTLGAVLGASLHTYGRSLYLPATEPALPIPGHSLPRAKSELYLAPAPARAATHTSP 380
QY 216 -----GRVGS----- 220
DB 381 LHRVSVDLPPLPOPALVSPGSPEDLPDLSLNGSGPELTGDMGAGADAPLSPDPLPSSP 440
QY 221 -----KTPAS-----PVYVQGGPVDAHLVLPREPVSHHTVVPD 254
DB 441 NALKTAOLSPFAEGPPEMPASSGFLPPHSAPEASTWGVDDLPLPLPSLPDQDVVTEG 500
QY 255 EDFVLVIALHSHLSEMFAPMGCAAGVNLFFYRAGVSRAMLRLFLMEKGRHMEY 314
DB 501 E-FVLMILILSHLCADLMAARHAFDDGVYHLCVRSIGISRAALLRLILAEHNHSL 559
QY 315 ECPYLVYVVAFRLEPPDKGVFAVDEGLMSEAVNGOVHNYWMSGCVPEPPSKRP 374
DB 560 GCPHIGVAAARFRLEPLTPKGLLTVDGLVEYGPIDQVHPGLATLTLTG----PAGQRP 615
QY 375 Q 375
DB 616 Q 616

RESULT 3
SPH2_HUMAN STANDARD: PRT; 654 AA.
AC O9NRAO; O9NRU7; O9HQ02; O9BRN1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sphingosine kinase 2 (EC 2.7.1.-) (SK 2) (SPK 2).
GN SPHK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
ON NCB1; TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
RA MEDLINE-20347850; Pubmed-10751414;
Liu H., Sugitara M., Nava V.E., Edsall L.C., Kono K., Poulton S.,

```

```

RA Milstien S., Kohama T., Spiegel S.;
RT "Molecular cloning and functional characterization of a novel
RT mammalian sphingosine kinase type 2 isoform.";
RL J. Biol. Chem. 275:19513-19520(2000).
(2)
RP SEQUENCE FROM N.A. (ISOFORM 2).
RP TISSUE-Brain;
RC MEDLINE-21154917; Pubmed-11230166;
RA Wiemann S., Wei B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansgore W., Boecker M., Bloecher H., Bauersachs S., Blum H.,
RA Mewes H.-W., Oltmannseder B., Obermaier K., Strack N.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
(3)
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE-Eye, and Lymph;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
(4)
RP SEQUENCE OF 1-354 FROM N.A. (ISOFORM 3).
RC TISSUE-Carcinoma;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohtsuka M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isgai T., Sugano S.;
RT "NEO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
CC - FUNCTION: Catalyzes the phosphorylation of sphingosine to form
CC sphingosine 1-phosphate (SPP), a lipid mediator with both intra-
CC and extracellular functions. Also acts on D-erythro-
CC dihydrosphingosine, D-erythro-sphingosine and L-threo-
CC dihydrosphingosine.
CC - CATALYTIC ACTIVITY: Sphingosine + ATP = sphingosine 1-phosphate +
CC ADP.
CC - ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; may be
CC produced by alternative splicing.
CC - SIMILARITY: BELONGS TO THE SPHINGOSINE KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: AF245447; AAF74124.1; -
DR EMBL: AL136701; CAB6636.1; -
DR EMBL: BC006161; AAH06161.1; -
DR EMBL: BC010671; AAH10671.1; -
DR EMBL: AK000599; BAA91280.1; -
DR MIM: 607092; -
DR Genew: HGNC:18859; SPHK2.
DR InterPro: IPR001206; DAGKc.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00781; DAGKc; 1.
DR ProDom: PD005043; DAG_kin_cat; 1.
DR SMART: SM00046; DAGKc; 1.
DR Transferrase: Kinase; ATP-binding.
FT CONFLICT 1 36 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPLIC 1 36 FEPLAGLDLILCSLLCRGGGHPDLISVTLASGRFSF
FT VARSPLIC 292 390 LSVANGTSDVDIOSERRALASARFTICGLATLTTR
FT GRSLTPATVPEASPTP -> PRESDSSSTSSACPLMTTA
FT RSCPRAAASMPGSCCLLPQALAGFSRPIQDVNCGGRIG
FT SLTCGATORTLPAPAREGGGSLPLKININVFICKKKK
FT (IN ISOFORM 3).
FT P -> S (IN REF. 2).
FT CONFLICT 49 49
SQ SEQUENCE 654 AA; 69217 MM; F73FCEC930DA50F CRC64;

Query Match 44.0%; Score 886.5; DB 1; Length 654;
Best Local Similarity 39.8%; Pred. No. 1.9e-69;

```


Db 160 ISYF-----TSALRTV-----SSAASFPMTLK---IDGFEI-----187
QY 250 TVPDEDFVLVALLHSHGSEMFAPMGRCAGVHILFYRAGVSRLMLRLFLAMEKG 309
Db 188 ----KEEYVMLVNGOIGINRIPLPDASTEDLDLICRN--TITLALRELMSEOG 241
QY 310 RHMEYECPLYVYVAVFRLEPKDKGFVAVDGLMWSAVAGVHPHYFMV 362
Db 242 -SIDRFGLSTVYVQASRIETIDPAKKA-DMDGEVYTHTPAVIQVLPHIDML 292

RESULT 5
Y036.SYNY3
ID Y036.SYNY3 STANDARD; PRT: 433 AA.
AC 055452;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein sl10036.
GN sl10036.
OS Synechocystis sp. (strain PCC 6803)
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneo T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
Sugiura M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
METHYLGLYOXAL SYNTHASE FAMILY.
CC -1- SIMILARITY: TO B.SOBITILIS BMRU AND TO E.COLI YEGS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D64006; BAA10795.1; -
DR HSP: P37066; B93
DR InterPro: IPR001206; DAGC.
DR InterPro: IPR003622; DAG_Kin_cat.
DR InterPro: IPR004362; DAG_Kin_cat.
DR InterPro: IPR004363; methylgl_synth.
DR Pfam: PF00781; DAGC; 1.
DR Pfam: PF02142; MGS; 1.
DR ProDom: PD005043; DAG_Kin_cat; 1.
DR ProDom: PD010761; Methylgl_synth; 1.
DR SMART: SM00046; DAGC; 1.
DR TIGRfams: TIGR00160; MGS; 1.
DR PROSITE: PS01335; METHYLGLYOXAL SYNTH; 1.
DR Hypothetical protein: Complete proteome.
FT DOMAIN 1 126 METHYLGLYOXAL SYNTHASE.
FT ACT_SITE 62 62
SQ SEQUENCE 433 AA; 45909 MW; BDB3D47B2990F6C5 CRC64;

Query March 5.0%; Score 100.5; DB 1; Length 433;
Best Local Similarity 26.8%; Pred. No. 0.34; Indels 23; Gaps 8;
Matches 40; Conservative 26; Mismatches 60;

QY 19 VLNRGCKG---KAIQVFRSHVQPLAEATISFTIMLTERRHARELVR-----SEEL 69
Db 129 LIPNVAGQGVNERLDLKEHLQSEI-NLKITFSAEVNTDQAKELVKRKIQANEQSD 187
QY 70 GRMDALVYMS-GDGLMHEVNGLMERPMEWTAIQKPLCSLPAGSGNALASL---NHAG 125
Db 188 GEGDSFIASGDGTGVSVAALV-----NIGI--PLGIIPRGTAAPSVALGIPTQIPG 240

QY 126 YEQVNEEDLTNCTLLICRRLSPMNILS 154
Db 241 ACOTINRGITRVVDALCNDI--PALLIA 267

RESULT 6
KDGL_HUMAN
ID KDGL_HUMAN STANDARD; PRT: 1065 AA.
AC 075912; Q9NZ49;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Diacylglycerol kinase, iota (EC 2.7.1.107) (Diacylceride kinase) (DGK-
iota) (DAG kinase iota).
GN DGKI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=99047655; PubMed=9830018;
RA Ding L., Traer E., McIntyre T.M., Zimmerman G.A., Prescott S.M.;
RT "The cloning and characterization of a novel human diacylglycerol
kinase, DGK-iota.";
RL J. Biol. Chem. 273:32746-32752(1998).
RN [2]
RP SEQUENCE OF 135-1065 FROM N.A., AND VARIANT PHE-153.
RX MEDLINE=20173854; PubMed=10706894;
RA Bowne S.J., Sullivan L.S., Ding L., Traer E., Prescott S.M.,
Birch D.G., Kennan A., Humphries P., Daiger S.P.;
RT "Evaluation of human diacylglycerol kinase iota, DGKI, a homolog of
Drosophila rfga, in inherited retinopathy mapping to 7q.";
RL Mol. Vision 6:6-9(2000).
CC -1- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol -> ADP + 1,2-
diacylglycerol 3-phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC DIACYLGLYCEROL KINASE
FAMILY.
CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
BINDING DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 MARCKS HOMOLOG REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF061936; AAC62010.1; -
DR EMBL: AF219939; AAF43006.1; -
DR EMBL: AF219907; AAF43006.1; JOINED.
DR EMBL: AF219908; AAF43006.1; JOINED.
DR EMBL: AF219909; AAF43006.1; JOINED.
DR EMBL: AF219910; AAF43006.1; JOINED.
DR EMBL: AF219911; AAF43006.1; JOINED.
DR EMBL: AF219912; AAF43006.1; JOINED.
DR EMBL: AF219913; AAF43006.1; JOINED.
DR EMBL: AF219914; AAF43006.1; JOINED.
DR EMBL: AF219915; AAF43006.1; JOINED.
DR EMBL: AF219916; AAF43006.1; JOINED.
DR EMBL: AF219917; AAF43006.1; JOINED.
DR EMBL: AF219918; AAF43006.1; JOINED.
DR EMBL: AF219919; AAF43006.1; JOINED.
DR EMBL: AF219920; AAF43006.1; JOINED.
DR EMBL: AF219921; AAF43006.1; JOINED.
DR EMBL: AF219922; AAF43006.1; JOINED.
DR EMBL: AF219923; AAF43006.1; JOINED.


```

CC -1- FUNCTION: UPON CELL STIMULATION CONVERTS THE SECOND MESSENGER
CC DIACYLGLYCEROL INTO PHOSPHATIDATE, INITIATING THE RESYNTHESIS
CC OF PHOSPHATIDYLINOSITOLS AND ATTENUATING PROTEIN KINASE C
CC ACTIVITY (BY SIMILARITY)
CC -1- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-
CC diacylglycerol 3-phosphate.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE NERVOUS SYSTEM AND MUSCLE.
CC -1- DEVELOPMENTAL STAGE: DGK IS TRANSCRIBED IN THE EMBRYONIC, PUPAL,
CC AND ADULT STAGES, WITH LITTLE EXPRESSION DURING THE LARVAL
CC STAGES. EXPRESSION IN LATE EMBRYOS IS SPECIFIC TO THE CENTRAL
CC NERVOUS SYSTEM AND HEAD.
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC DIACYLGLYCEROL KINASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D11120; BAA01894.1; -
DR EMBL: AE003839; AAF59180.1; -
DR EMBL: X67335; CAA47750.1; -
DR PIR: S25099; S25099.
DR PIR: S28329; S28329.
DR PIR: A46140; A46140.
DR FlyBase: FBgn0004568; Dgk.
DR InterPro: IPR000756; DAGK.
DR InterPro: IPR001206; DAGKC.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00609; DAGKa; 1.
DR Pfam: PF00781; DAGKc; 1.
DR ProDom: PD002939; DAGKa; 1.
DR ProDom: PD005043; DAG_kin_cat; 1.
DR SMART: SM00045; DAGKa; 1.
DR SMART: SM00046; DAGKc; 1.
DR KJ Transferase: Kinase; 1.
FT DOMAIN 311 372 THR-RICH.
FT DOMAIN 401 405 POLY-ALA.
FT DOMAIN 430 453 GLN-RICH.
FT DOMAIN 534 539 POLY-GLN.
FT DOMAIN 130 256 CATALYTIC-A (POTENTIAL).
FT DOMAIN 564 755 CATALYTIC-B (POTENTIAL).
FT CONFLICT 374 374 M -> T (IN REF. 1).
FT CONFLICT 444 444 Q -> QQQQQ (IN REF. 3).
SQ SEQUENCE 791 AA; 87297 MW; 9DFD00E280FD64A6 CRC64;

Query Match 4.6%; Score 93.5; DB 1; Length 791;
Best Local Similarity 23.1%; Pred. No. 3.1;
Matches 49; Conservative 34; Mismatches 86; Indels 43; Gaps 9;

OY 15 CRVLVLLNPGKGGKKAQLQFRSHVQPLLAELAEISFTMLTERRNH-----ARELVASE 67
DB 131 CPLLIVFVNPKSGRGQGRDILR-----KFOYMLNPROVVDLSKGGKRECLTJFK 178
OY 68 ELGRWDALVWNSGGLMHEVYVNGILMERPD-WETALQKFLCSIPAGSGNALASLHNYAGY 126
DB 179 DLPRF-RVICGGDGG---TVGWVLEAMDSIELASQPAIGVPLCTGNDLARCLRWGGGY 233
OY 127 EOVNEDLLTN-----CTLLRCRLISPMNL-----LSLHTASGLRFSVLSLAWGF 173
DB 234 EGENIPKLMDFRRASTYMLDRWSIEVNTNTHPSDMRKRVILHS-----NMOKVIELSOV 289
OY 174 IADVLESEKYRRLGEMFTLTGTLRLAALRT 205
DB 290 VYDKSL-MERFEELIOROSKOVATSWGTAAST 320

RESULT 8
DNLJ_RHOMR STANDARD; PRT; 712 AA.
ID DNLJ_RHOMR

```

```

AC P49421;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA ligase (EC 6.5.1.2) (Polydeoxyribonucleotide synthase [NAD+]).
GN LIGA OR LIG.
OS Rhotothermus marinus (Rhotothermus obamensis).
OC Bacteria; Bacteroidetes; Sphingobacteriales;
OC Chromatiales; Rhotothermus.
OX NCBI_TaxID=29549;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R-21;
RX MEDLINE=95369716; PubMed=7642120;
RA Kristianson J.K., Eggertson G., Palstoft A.;
RT "Cloning and sequence analysis of the DNA ligase-encoding gene of
RT Rhotothermus marinus, and overproduction, purification and
RT characterization of two thermophilic DNA ligases.";
RL Gene 161:1-6(1995).
CC -1- FUNCTION: THIS PROTEIN CATALYZES THE FORMATION OF PHOSPHODIESTER
CC LINKAGES BETWEEN 5'-PHOSPHORYL AND 3'-HYDROXYL GROUPS IN DOUBLE-
CC STRANDED DNA USING NAD AS A COENZYME AND AS THE ENERGY SOURCE FOR
CC THE REACTION. IT IS ESSENTIAL FOR DNA REPLICATION AND REPAIR OF
CC DAMAGED DNA. THIS ENZYME IS THERMOSTABLE BEING ACTIVE AT 5-75
CC DEGREES CELSIUS WITH APPARENT OPTIMAL ACTIVITY ABOVE 55 DEGREES
CC CELSIUS.
CC -1- CATALYTIC ACTIVITY: NAD(+) + [deoxyribonucleotide](N) +
CC [deoxyribonucleotide](M) = AMP + nicotinamide nucleotide +
CC [deoxyribonucleotide](N+M).
CC -1- MISCELLANEOUS: THIS ENZYME IS THERMOSTABLE.
CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT DNA LIGASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U10483; AAA93198.1; -
DR HSSP: 087703; 1B04.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR004150; DNA_ligase_OB.
DR InterPro: IPR001679; DNA_ligase.
DR InterPro: IPR003265; Endo_3c.
DR InterPro: IPR000445; HNH.
DR InterPro: IPR004149; Znf_DNA_ligase_C4.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00633; HNH; 1.
DR Pfam: PF01653; DNA_ligase_N; 1.
DR Pfam: PF03119; DNA_ligase_ZBD; 1.
DR Pfam: PF03120; DNA_ligase_OB; 1.
DR ProDom: PD003944; DNAligase; 1.
DR SMART: SM00292; BRCT; 1.
DR SMART: SM00478; ENDO3c; 1.
DR SMART: SM00278; HNH; 3.
DR SMART: SM00532; LIGANC; 1.
DR TIGRFAMs: TIGR00575; dnlj; 1.
DR PROSITE: PS01072; BRCT; 1.
DR PROSITE: PS01055; DNA_LIGASE_N1; 1.
DR PROSITE: PS01056; DNA_LIGASE_N2; 1.
DR Ligase; DNA repair; DNA replication; NAD.
FT DOMAIN 633 712 BRCT.
FT BINDING 153 153 AMP (BY SIMILARITY).
SQ SEQUENCE 712 AA; 79487 MW; 893AD3A78F77EFC1 CRC64;

Query Match 4.6%; Score 93; DB 1; Length 712;
Best Local Similarity 21.8%; Pred. No. 2.9;
Matches 63; Conservative 37; Mismatches 109; Indels 80; Gaps 11;

```

```

OY 59 HARELVSPELGRMDALVMSGDLMEHYVNGLM-----ERPMETAIOKPLCS----- 107
DB 401 HNEEDYRSRDIRIGDLVYIRAGDVIPOVVRPVYERNGNERP-WRRPEPCSGSOLVR 459
OY 108 LPAGSGNMLAAS-----LHNYAGYEQYTNEDLLNCLLRCR-LTSPM----- 150
DB 460 LRGADYCVASDPCDPAQVRLLEHFAGRANDMEGSGVAVAROLABESGLVPLSDLYRLK 519
OY 151 --NLSTHTASGLRFSVLSLAFADVLDSEKRYRLGEMKFTLGTFLRLAALPTKRG 208
DB 520 LBDLKLBEFAFTRARNLRA-----TEASKQRLSLRLFLGLG-----IRHVGR 563
OY 209 RLAVLPVGVSGKTPASPVVQQ-----GPVDAHLVP----- 240
DB 564 TTAELLYORFASIDELAAVITDEIALLEGVGPITAESIANMFREVDNRRLIEELKEIGVN 623
OY 241 ---LEBPVSHMTVVPDEFLVYLLALH--SHLGSEMFAPMCRCAAGV 284
DB 624 TORLPEAPAAESPVRKGTFTVLGALPHLTRKEABELIKRAGGVAVASSV 672

RESULT 9
ID YSM3_CAEEL STANDARD: PRT: 439 AA.
AC 010123:
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein f52c9.3 in chromosome III.
GN f52c9.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitidae; Pelodermidae; Caenorhabdilitis.
OX NCBI_TaxID=6239;
RN 11
RC SEQUENCE FROM N.A.
RA STRAIN-Bristol N2;
RA Favalio T.;
RA Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN 12
RP REVISIONS.
RA Waterston R.;
RA Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U39850; AAA81060.2;
DR Wormpep: f52c9.3; CE29806;
DR Interpro: IPR001206; DACC;
DR Interpro: IPR003622; DAG_Kin_cat;
DR Prodom: PD005043; DAG_Kin_cat; I;
DR SMART: SM00046; DAGKc; I;
KW Hypothetical protein;
SQ SEQUENCE 439 AA: 49974 MW: 779850DCG9662DBB CRC64:

Query Match 4.6%; Score 92.5; DB 1; Length 439;
Best Local Similarity 26.5%; Pred. No. 1.7;
Matches 27; Conservative 14; Mismatches 58; Indels 3; Gaps 1;

```

```

RESULT 10
ID GV22_HA1N1 STANDARD: PRT: 492 AA.
AC 09HHT12:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GVPD protein 2.
GN GVPD2 OR VNG6240G.
OS Halobacterium sp. (strain NRC-1).
OC Plasmid pNRC200.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shroana J.,
RA Swartzell S., Welt D., Hall J., Dahl T.A., Welt V., Goo Y.A.,
RA Leitbauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenberger T.A., Peck R.F., Pohlischer M., Spudich J.I., Jung K.-H.,
RA Alem M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarum S.;
RA "Genome sequence of Halobacterium species NRC-1";
RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- FUNCTION: MAY PLAY A ROLE IN AN ENERGY REQUIRING PROCESS SUCH AS
CC ASSEMBLY OF GAS VESTICLES IN ADDITION TO A POSSIBLE STRUCTURAL OR
CC REGULATORY FUNCTION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF005156; AAG20894.1;
DR Gas vesicle; ATP-binding; Plasmid; Complete proteome.
KW NP_BIND;
FT 39 46
SQ SEQUENCE 492 AA: 53928 MW: 655965FDIC36331B CRC64:

Query Match 4.6%; Score 92.5; DB 1; Length 492;
Best Local Similarity 24.3%; Pred. No. 2;
Matches 89; Conservative 36; Mismatches 146; Indels 95; Gaps 16;

```


